VSphere Metro Storage Cluster con ONTAP

Enterprise applications

NetApp
May 03, 2024

Sommario

VSphere Metro Storage Cluster con ONTAP

VSphere Metro Storage Cluster con ONTAP

L'hypervisor vSphere leader del settore di VMware può essere implementato come cluster stretched indicato come vSphere Metro Storage Cluster (vMSC).

Le soluzioni vMSC sono supportate sia con NetApp® MetroCluster™ che con SnapMirror Active Sync (precedentemente noto come SnapMirror Business Continuity o SMBC) e forniscono una business continuity avanzata se uno o più domini di errore subiscono un'interruzione totale. La resilienza alle diverse modalità di errore dipende dalle opzioni di configurazione scelte.

Soluzioni di disponibilità continua per ambienti vSphere

L'architettura ONTAP è una piattaforma di storage flessibile e scalabile che fornisce servizi SAN (FCP, iSCSI e NVMe-of) e NAS (NFS v3 e v4,1) per datastore. I sistemi storage NetApp AFF, ASA e FAS utilizzano il sistema operativo ONTAP per offrire protocolli aggiuntivi per l'accesso allo storage guest, come S3 e SMB/CIFS.

NetApp MetroCluster utilizza la funzione di ha (failover del controller o CFO) di NetApp per la protezione dai guasti dei controller. Include inoltre la tecnologia SyncMirror locale, il failover cluster in caso di disastro (failover controller on-demand o CFOD), la ridondanza hardware e la separazione geografica per ottenere livelli elevati di disponibilità. SyncMirror esegue il mirroring sincrono dei dati tra le due metà della configurazione MetroCluster scrivendo i dati su due plessi: Il plesso locale (sullo shelf locale) fornendo attivamente i dati e il plesso remoto (sullo shelf remoto) normalmente non fornendo i dati. La ridondanza hardware viene implementata per tutti i componenti MetroCluster, come controller, storage, cavi, switch (utilizzati con Fabric MetroCluster) e adattatori.

La sincronizzazione attiva di NetApp SnapMirror fornisce una protezione granulare dei datastore con protocolli SAN FCP e iSCSI, permettendoti di proteggere in modo selettivo solo i carichi di lavoro ad alta priorità. Offre l'accesso Active-Active ai siti locali e remoti, a differenza di NetApp MetroCluster, che è una soluzione Active-standby. Attualmente, la sincronizzazione attiva è una soluzione asimmetrica in cui un lato è preferito rispetto all'altro, fornendo prestazioni migliori. Ciò si ottiene utilizzando la funzionalità ALUA (Asymmetric Logical Unit Access) che informa automaticamente l'host ESXi, quali controller preferire. Tuttavia, NetApp ha annunciato che la sincronizzazione attiva presto abiliterà l'accesso completamente simmetrico.

Per creare un cluster VMware ha/DRS su due siti, gli host ESXi vengono utilizzati e gestiti da un'appliance vCenter Server (VCSA). Le reti di gestione vSphere, vMotion® e delle macchine virtuali sono collegate tramite una rete ridondante tra i due siti. VCenter Server che gestisce il cluster ha/DRS può connettersi agli host ESXi in entrambi i siti e deve essere configurato utilizzando vCenter ha.

Fare riferimento a. "Come creare e configurare i cluster nel client vSphere" Per configurare vCenter ha.

Fare riferimento anche alla sezione "Procedure consigliate per VMware vSphere Metro Storage Cluster".

Che cos'è vSphere Metro Storage Cluster?

VSphere Metro Storage Cluster (vMSC) è una configurazione certificata che protegge le macchine virtuali (VM) e i container dai guasti. Ciò si ottiene utilizzando concetti di storage estesi insieme ai cluster di host ESXi, distribuiti in diversi domini di errore come rack, edifici, campus o persino città. Le tecnologie di storage Active Sync di NetApp MetroCluster e SnapMirror vengono utilizzate per fornire ai cluster host una protezione rispettivamente con RPO=0 o near RPO=0. La configurazione vMSC è progettata per garantire che i dati siano sempre disponibili, anche in caso di errore di un "sito" fisico o logico completo. Un dispositivo di storage che fa

parte della configurazione vMSC deve essere certificato dopo aver superato un processo di certificazione vMSC di successo. Tutti i dispositivi di archiviazione supportati sono disponibili nella "Guida alla compatibilità dello storage VMware".

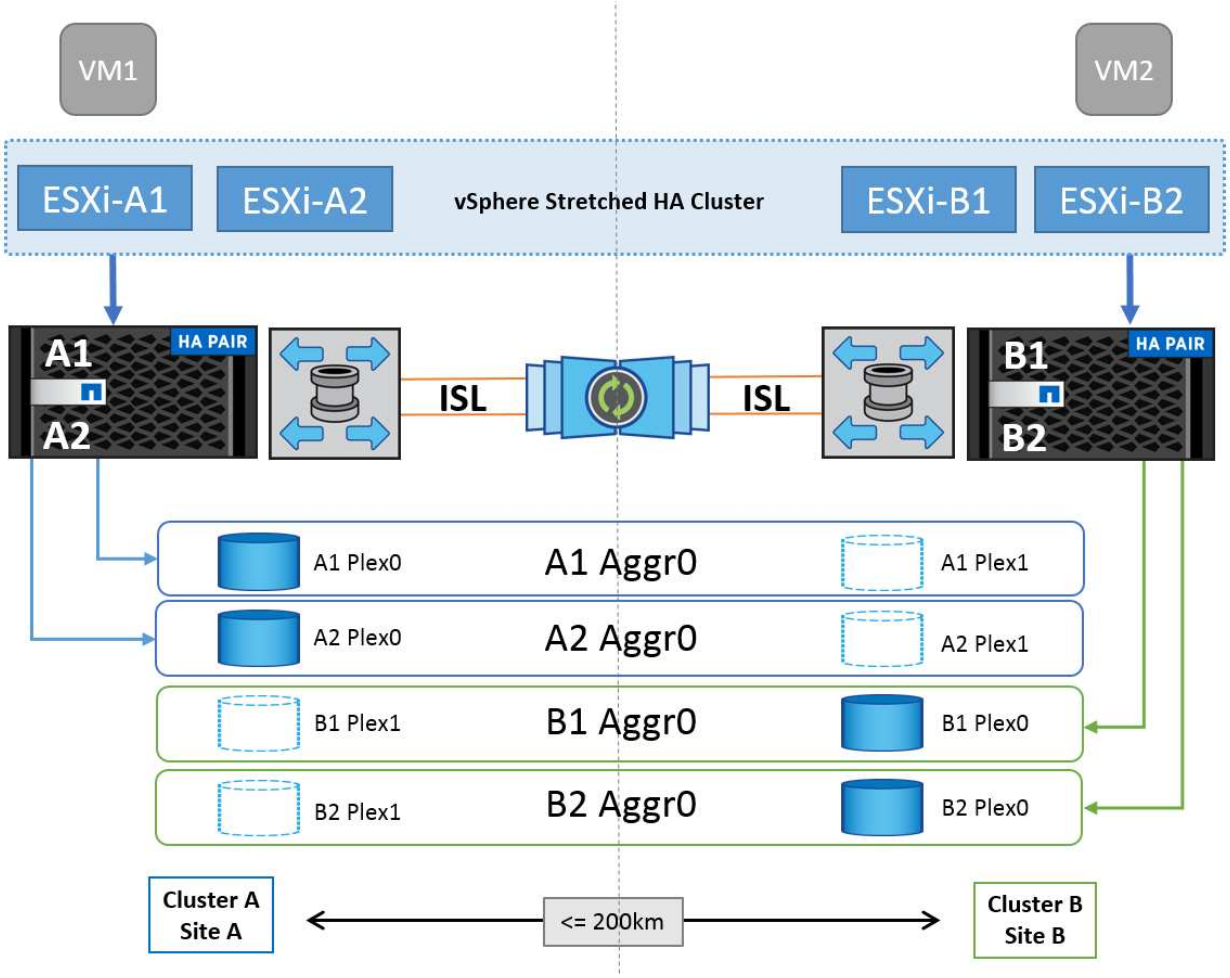
Per ulteriori informazioni sulle linee guida di progettazione per vSphere Metro Storage Cluster, consultare la seguente documentazione:

- "Supporto di VMware vSphere con NetApp MetroCluster"
- "Supporto di VMware vSphere con business continuity di NetApp SnapMirror" (Adesso noto come SnapMirror Active Sync)

A seconda delle considerazioni sulla latenza, NetApp MetroCluster può essere implementato in due diverse configurazioni da utilizzare con vSphere:

- Stretch MetroCluster
- Fabric MetroCluster

Di seguito viene illustrato uno schema topologico di alto livello di Stretch MetroCluster.

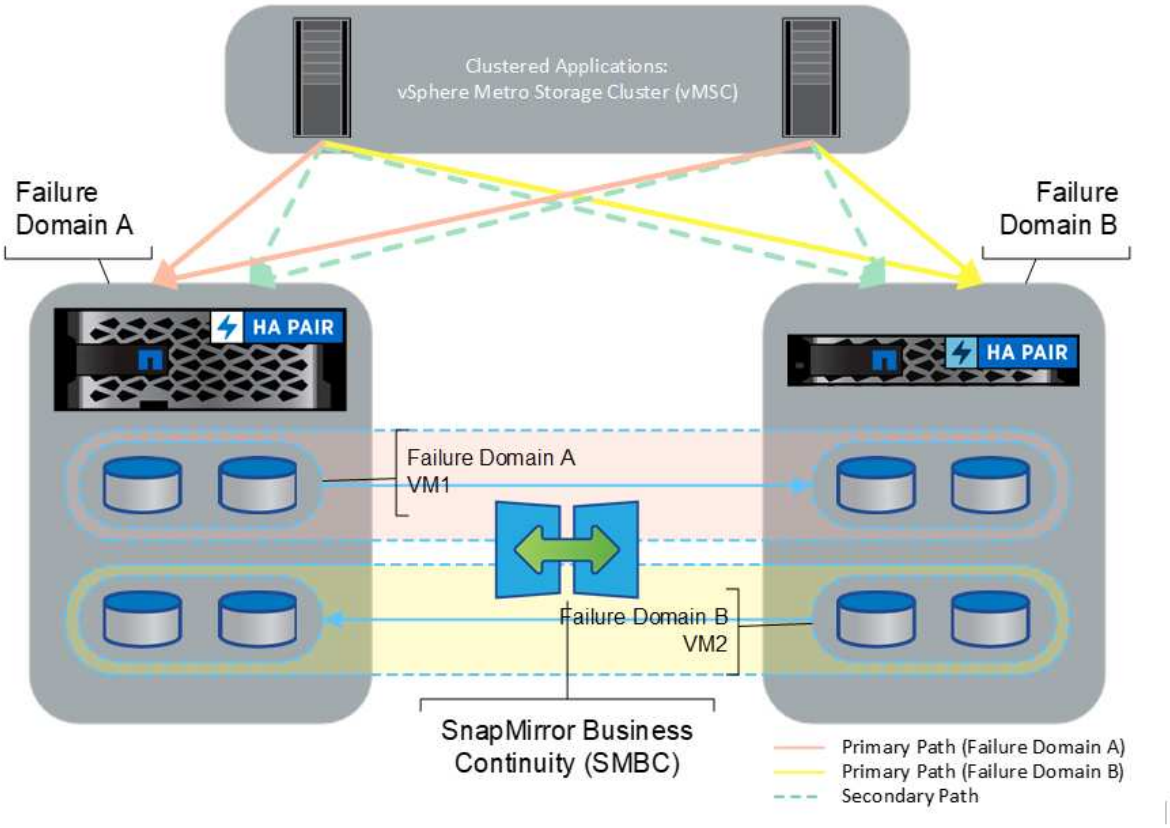


Fare riferimento a. "Documentazione MetroCluster" Per informazioni specifiche sulla progettazione e la distribuzione di MetroCluster.

SnapMirror Active Sync può anche essere implementato in due modi diversi.

- Asimmetrico

- Simmetrico (anteprima privata in ONTAP 9.14.1)



Fare riferimento a. "Documenti NetApp" Per informazioni specifiche sulla progettazione e la distribuzione per la sincronizzazione attiva di SnapMirror.

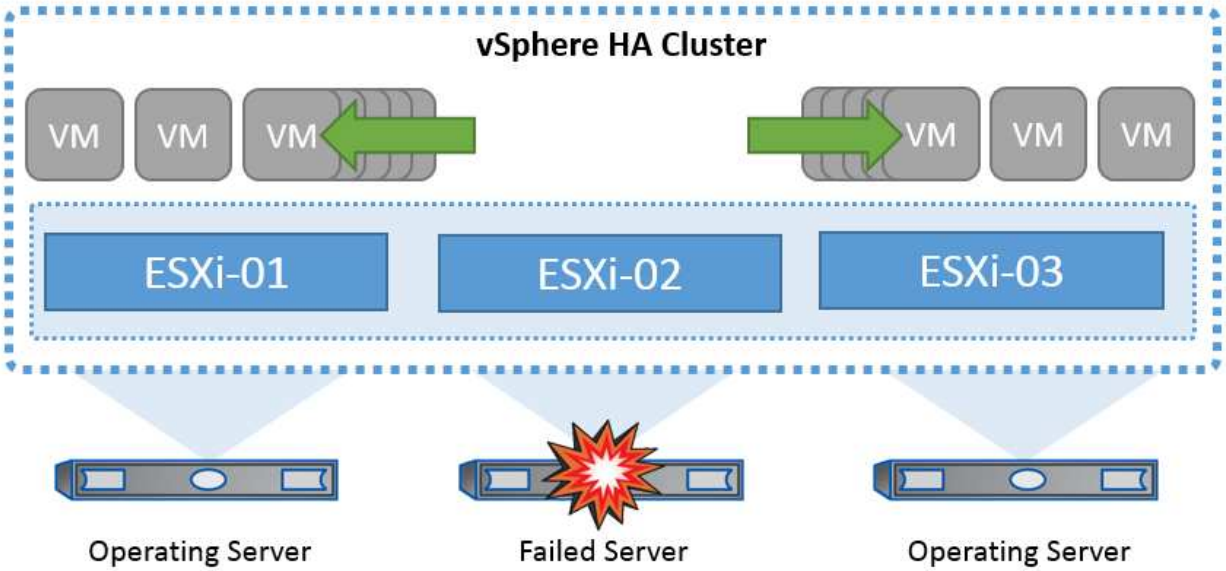
Panoramica della soluzione VMware vSphere

VCenter Server Appliance (VCSA) è un potente sistema di gestione centralizzato e un singolo pannello di controllo per vSphere che consente agli amministratori di utilizzare in modo efficace i cluster ESXi. Agevola le funzioni chiave come provisioning delle macchine virtuali, funzionamento di vMotion, alta disponibilità (ha), Distributed Resource Scheduler (DRS), Tanzu Kubernetes Grid e altro ancora. Si tratta di un componente essenziale negli ambienti cloud VMware e deve essere progettato tenendo presente la disponibilità del servizio.

Alta disponibilità vSphere

La tecnologia cluster di VMware raggruppa i server ESXi in pool di risorse condivise per le macchine virtuali e offre vSphere High Availability (ha). VSphere ha offre alta disponibilità e facile da utilizzare per le applicazioni eseguite su macchine virtuali. Quando la funzionalità ha è abilitata sul cluster, ogni server ESXi mantiene la comunicazione con altri host in modo che, se un host ESXi non risponde o si isola, il cluster di ha può negoziare il recovery delle macchine virtuali in esecuzione sull'host ESXi tra gli host sopravvissuti nel cluster. In caso di errore del sistema operativo guest, vSphere ha riavvia la macchina virtuale interessata sullo stesso server fisico. VSphere ha consente di ridurre i downtime pianificati, prevenire i downtime non pianificati e eseguire un rapido ripristino in caso di interruzioni.

Cluster vSphere ha in grado di ripristinare le VM dal server guasto.



È importante comprendere che VMware vSphere non conosce NetApp MetroCluster o SnapMirror Active Sync e vede tutti gli host ESXi nel cluster vSphere come host idonei per le operazioni del cluster ha in base alle configurazioni di affinità dei gruppi VM e host.

Rilevamento errori host

Non appena viene creato il cluster ha, tutti gli host nel cluster partecipano alle elezioni e uno degli host diventa un master. Ogni slave esegue heartbeat di rete al master, e il master a sua volta esegue heartbeat di rete su tutti gli host slave. L'host master di un cluster vSphere ha è responsabile del rilevamento del guasto degli host slave.

A seconda del tipo di errore rilevato, potrebbe essere necessario eseguire il failover delle macchine virtuali in esecuzione sugli host.

In un cluster vSphere ha, vengono rilevati tre tipi di errore dell'host:

- Errore - Un host smette di funzionare.
- Isolamento - Un host diventa isolato dalla rete.
- Partizione - Un host perde la connettività di rete con l'host master.

L'host master monitora gli host slave nel cluster. Questa comunicazione viene fatta attraverso lo scambio di heartbeat di rete ogni secondo. Quando l'host master smette di ricevere questi heartbeat da un host slave, controlla la liveness dell'host prima di dichiarare che l'host non è riuscito. Il controllo liveness che l'ospite principale effettua è di determinare se l'ospite secondario sta scambiando i heartbeat con uno dei datastore. Inoltre, l'host master verifica se l'host risponde ai ping ICMP inviati ai propri indirizzi IP di gestione per rilevare se è semplicemente isolato dal suo nodo master o completamente isolato dalla rete. Per farlo, eseguire il ping del gateway predefinito. È possibile specificare manualmente uno o più indirizzi di isolamento per migliorare l'affidabilità della convalida dell'isolamento.

Best practice

NetApp consiglia di specificare un minimo di due indirizzi di isolamento aggiuntivi e che ciascuno di questi indirizzi sia locale al sito. Ciò migliorerà l'affidabilità della convalida dell'isolamento.

Risposta di isolamento dell'host

Risposta di isolamento è un'impostazione in vSphere ha che determina l'azione attivata sulle macchine virtuali quando un host in un cluster vSphere ha perde le connessioni di rete di gestione ma continua a essere eseguito. Sono disponibili tre opzioni per questa impostazione: "Disabilitato", "Arresta e riavvia le macchine virtuali" e "Spegni e riavvia le macchine virtuali".

Lo "spegnimento" è migliore dello "spegnimento", che non svuota le modifiche più recenti al disco o esegue il commit delle transazioni. Se le macchine virtuali non si sono arrestate entro 300 secondi, vengono spente. Per modificare il tempo di attesa, utilizzare l'opzione avanzata das.isolationshutdowntimeout.

Prima che ha avvii la risposta di isolamento, verifica prima se l'agente master ha vSphere è proprietario del datastore che contiene i file di configurazione della VM. In caso contrario, l'host non attiverà la risposta di isolamento, poiché non vi è alcun master per riavviare le VM. L'host controllerà periodicamente lo stato del datastore per determinare se viene richiesto da un agente vSphere ha che detiene il ruolo master.

Best practice

NetApp consiglia di impostare la risposta di isolamento dell'host su Disabilitato.

Una condizione split-brain può verificarsi se un host viene isolato o partizionato dall'host master vSphere ha e il master non è in grado di comunicare tramite datastore heartbeat o tramite ping. Il master dichiara l'host isolato inattivo e riavvia le macchine virtuali su altri host nel cluster. Esiste ora una condizione split-brain perché esistono due istanze della macchina virtuale in esecuzione, una sola delle quali è in grado di leggere o scrivere i dischi virtuali. Le condizioni split-brain possono ora essere evitate configurando VMCP (VM Component Protection).

Protezione dei componenti VM (VMCP)

Uno dei miglioramenti delle funzionalità di vSphere 6, relativi all'ha, è VMCP. VMCP fornisce una protezione avanzata da APD (All Path Down) e PDL (Permanent Device Loss) per lo storage a blocchi (FC, iSCSI, FCoE) e a file (NFS).

Perdita permanente del dispositivo (PDL)

PDL è una condizione che si verifica quando un dispositivo di memorizzazione si guasta in modo permanente o viene rimosso amministrativamente e non deve essere restituito. L'array di storage NetApp invia un codice di rilevamento SCSI a ESXi dichiarando che il dispositivo è perso in modo permanente. Nella sezione Condizioni di guasto e Risposta VM di vSphere ha, è possibile configurare la risposta che deve essere dopo il rilevamento di una condizione PDL.

Best practice

NetApp consiglia di impostare "Risposta per datastore con PDL" su "**Spegni e riavvia VM**". Quando viene rilevata questa condizione, una VM viene riavviata istantaneamente su un host integro all'interno del cluster vSphere ha.

Tutti i percorsi verso il basso (APD)

APD è una condizione che si verifica quando un dispositivo di archiviazione diventa inaccessibile all'host e non sono disponibili percorsi all'array. ESXi considera questo un problema temporaneo con il dispositivo e si aspetta che diventi nuovamente disponibile.

Quando viene rilevata una condizione APD, viene avviato un timer. Dopo 140 secondi, la condizione APD viene dichiarata ufficialmente e il dispositivo viene contrassegnato come timeout APD. Una volta trascorsi i 140

secondi, ha inizia il conteggio dei minuti specificati nell'APD Delay for VM failover. Una volta trascorso il tempo specificato, ha riavvia le macchine virtuali interessate. È possibile configurare VMCP in modo che risponda in modo diverso, se lo si desidera (Disattivato, Eventi problema o Spegni e riavvia le macchine virtuali).

Best practice

NetApp consiglia di configurare "Risposta per datastore con APD" su "**Spegni e riavvia le VM (conservative)**".

Conservative si riferisce alla probabilità che ha sia in grado di riavviare le VM. Quando è impostata su Conservative, ha riavvia la VM interessata dall'APD solo se sa che un altro host può riavviarla. In caso di problemi aggressivi, ha tenterà di riavviare la macchina virtuale anche se non conosce lo stato degli altri host. Ciò può comportare il mancato riavvio delle VM se non vi è alcun host con accesso al datastore su cui si trova.

Se lo stato APD viene risolto e l'accesso allo storage viene ripristinato prima del termine del timeout, l'ha non riavvia inutilmente la macchina virtuale a meno che non sia stata configurata esplicitamente. Se si desidera una risposta anche quando l'ambiente è stato ripristinato dalla condizione APD, è necessario configurare la risposta per il ripristino APD dopo il timeout APD in modo da ripristinare le VM.

Best practice

NetApp consiglia di configurare la risposta per il ripristino APD dopo il timeout APD su Disabilitato.

Implementazione VMware DRS per NetApp MetroCluster

VMware DRS è una funzionalità che aggrega le risorse host in un cluster e viene utilizzata principalmente per il bilanciamento del carico all'interno di un cluster in un'infrastruttura virtuale. VMware DRS calcola principalmente le risorse di CPU e memoria per eseguire il bilanciamento del carico in un cluster. Poiché vSphere non è consapevole del clustering allungato, considera tutti gli host in entrambi i siti durante il bilanciamento del carico. Per evitare il traffico tra siti, NetApp consiglia di configurare le regole di affinità DRS per gestire una separazione logica delle VM. In questo modo si garantisce che, a meno che non si verifichi un errore completo del sito, ha e DRS utilizzino solo host locali.

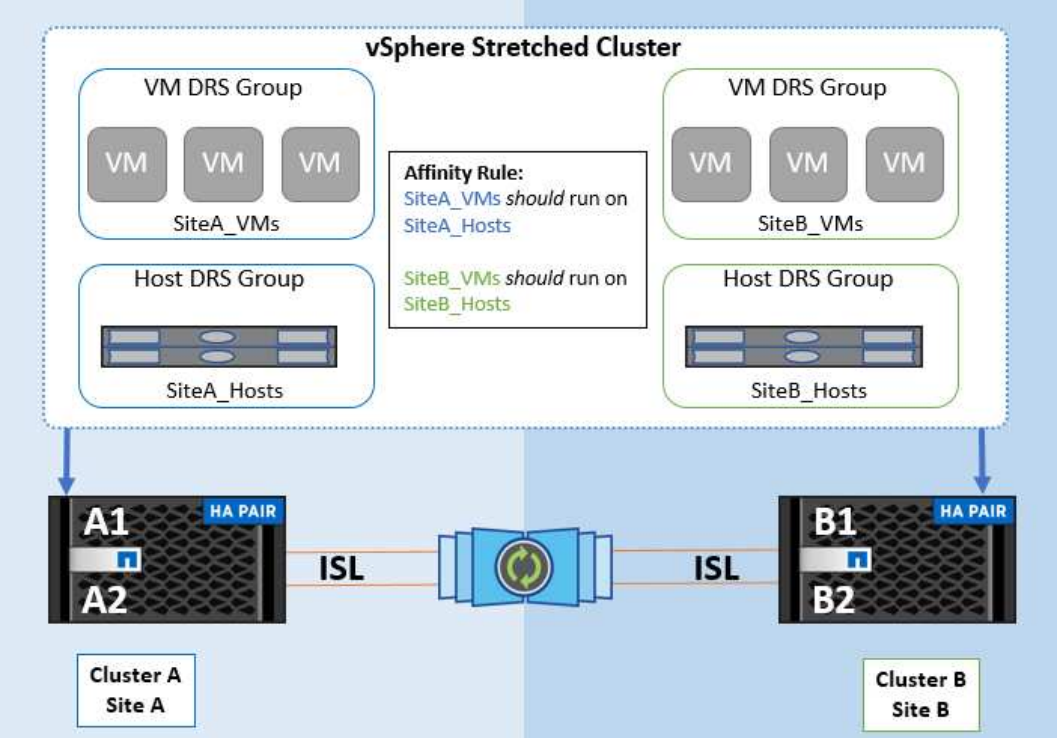
Se si crea una regola di affinità DRS per il cluster, è possibile specificare in che modo vSphere applica tale regola durante il failover di una macchina virtuale.

Esistono due tipi di regole che è possibile specificare il comportamento di failover di vSphere ha:

- Le regole di anti-affinità delle macchine virtuali costringono le macchine virtuali specificate a rimanere separate durante le azioni di failover.
- Le regole di affinità degli host VM collocano macchine virtuali specifiche su un host specifico o su un membro di un gruppo definito di host durante le azioni di failover.

Utilizzando le regole di affinità degli host delle macchine virtuali in VMware DRS, si può avere una separazione logica tra il sito A e il sito B in modo che la macchina virtuale venga eseguita sull'host nello stesso sito dell'array configurato come controller di lettura/scrittura principale per un determinato datastore. Inoltre, le regole di affinità degli host delle macchine virtuali consentono alle macchine virtuali di rimanere locali rispetto allo storage, il che a sua volta determina la connessione della macchina virtuale in caso di errori di rete tra i siti.

Di seguito è riportato un esempio di gruppi di host VM e regole di affinità.



Best practice

NetApp consiglia di implementare le regole "should" invece di quelle "must", in quanto vengono violate da vSphere ha in caso di errore. L'utilizzo di regole "must" può potenzialmente causare interruzioni del servizio.

La disponibilità dei servizi dovrebbe sempre prevalere sulle prestazioni. Nello scenario in cui si verifica un guasto di un data center completo, le regole "must" devono scegliere gli host dal gruppo di affinità degli host VM e, quando il data center non è disponibile, le macchine virtuali non verranno riavviate.

Implementazione di VMware Storage DRS con NetApp MetroCluster

La funzione VMware Storage DRS consente l'aggregazione di datastore in una singola unità e bilancia i dischi della macchina virtuale quando vengono superate le soglie di controllo i/o di storage.

Il controllo i/o dello storage è abilitato per impostazione predefinita sui cluster DRS abilitati per Storage DRS. Il controllo i/o dello storage consente a un amministratore di controllare la quantità di i/o dello storage allocata alle macchine virtuali nei periodi di congestione dell'i/o e di conseguenza le macchine virtuali più importanti possono preferire le macchine virtuali meno importanti per l'allocazione delle risorse i/O.

Storage DRS utilizza Storage vMotion per migrare le macchine virtuali in datastore diversi all'interno di un cluster di datastore. In un ambiente NetApp MetroCluster, la migrazione di una macchina virtuale deve essere controllata all'interno dei datastore di quel sito. Ad esempio, la macchina virtuale A, in esecuzione su un host nel sito A, dovrebbe idealmente migrare all'interno dei datastore della SVM nel sito A. In caso contrario, la macchina virtuale continuerà a funzionare ma con prestazioni ridotte, poiché la lettura/scrittura del disco virtuale avverrà dal sito B attraverso collegamenti tra siti.

Best practice

NetApp consiglia di creare cluster di datastore in relazione all'affinità con i siti storage. In altre parole, i datastore con affinità con i siti per il sito A non devono essere mescolati con i cluster di datastore con datastore con affinità con i siti per il sito B.

Ogni volta che viene eseguito il provisioning o la migrazione di una macchina virtuale mediante Storage vMotion, NetApp consiglia di aggiornare manualmente tutte le regole VMware DRS specifiche di tali macchine virtuali. In questo modo, si verificherà l'affinità della macchina virtuale a livello di sito per host e datastore, riducendo così l'overhead di rete e storage.

Linee guida per la progettazione e l'implementazione di vMSC

Questo documento delinea le linee guida di progettazione e implementazione per vMSC con i sistemi di storage ONTAP.

Configurazione dello storage NetApp

Le istruzioni per l'installazione di NetApp MetroCluster (definite configurazione MCC) sono disponibili all'indirizzo "Documentazione MetroCluster". Le istruzioni per la sincronizzazione attiva di SnapMirror sono disponibili all'indirizzo "Panoramica di SnapMirror Business Continuity".

Una volta configurato MetroCluster, gestirlo è come gestire un ambiente ONTAP tradizionale. Puoi configurare Storage Virtual Machine (SVM) utilizzando vari strumenti come l'interfaccia a riga di comando (CLI), System Manager o Ansible. Una volta configurate le SVM, occorre creare nel cluster interfacce logiche (LIF), volumi e LUN (Logical Unit Number) da utilizzare per le normali operazioni. Questi oggetti verranno replicati automaticamente sull'altro cluster utilizzando la rete di peering del cluster.

Se non utilizzi MetroCluster, puoi usare SnapMirror Active Sync che offre protezione granulare dei datastore e accesso Active-Active su diversi cluster ONTAP in diversi domini di errore. SnapMirror Active Sync utilizza gruppi di coerenza per garantire la coerenza dell'ordine di scrittura in uno o più datastore. Puoi creare più gruppi di coerenza in base ai requisiti di applicazioni e datastore. I gruppi di coerenza sono particolarmente utili per le applicazioni che richiedono la sincronizzazione dei dati tra datastore multipli. La sincronizzazione attiva di SnapMirror supporta inoltre RDM (Raw Device Mapping) e storage connesso al guest con initiator iSCSI in-guest. Per ulteriori informazioni sui gruppi di coerenza, visitare il sito Web all'indirizzo "Panoramica dei gruppi di coerenza".

Esiste una certa differenza nella gestione di una configurazione vMSC con sincronizzazione attiva SnapMirror rispetto a una MetroCluster. In primo luogo, si tratta DI una configurazione solo SAN, ma non è possibile proteggere datastore NFS con la sincronizzazione attiva di SnapMirror. In secondo luogo, è necessario mappare entrambe le copie delle LUN agli host ESXi per accedere ai datastore replicati in entrambi i domini di errore.

VMware vSphere ha

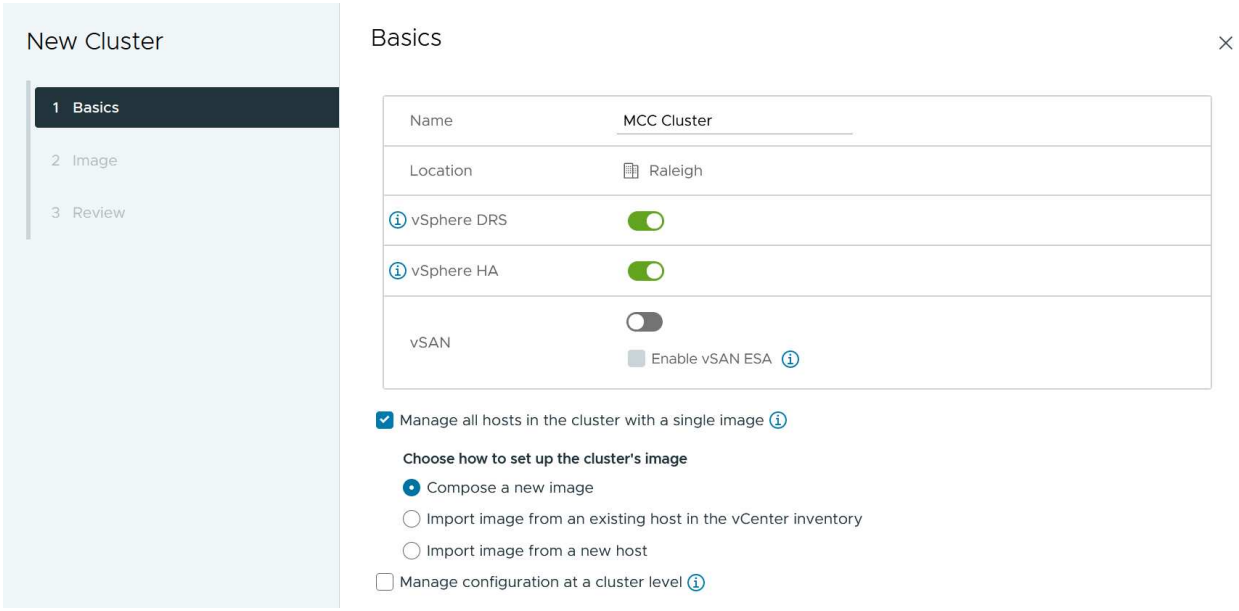
Creare un cluster vSphere ha

La creazione di un cluster vSphere ha è un processo in più fasi documentato all'indirizzo "Come creare e configurare i cluster nel client vSphere su docs.vmware.com". In poche parole, devi prima creare un cluster vuoto, quindi, utilizzando vCenter, devi aggiungere host e specificare l'ha vSphere del cluster e le altre impostazioni.

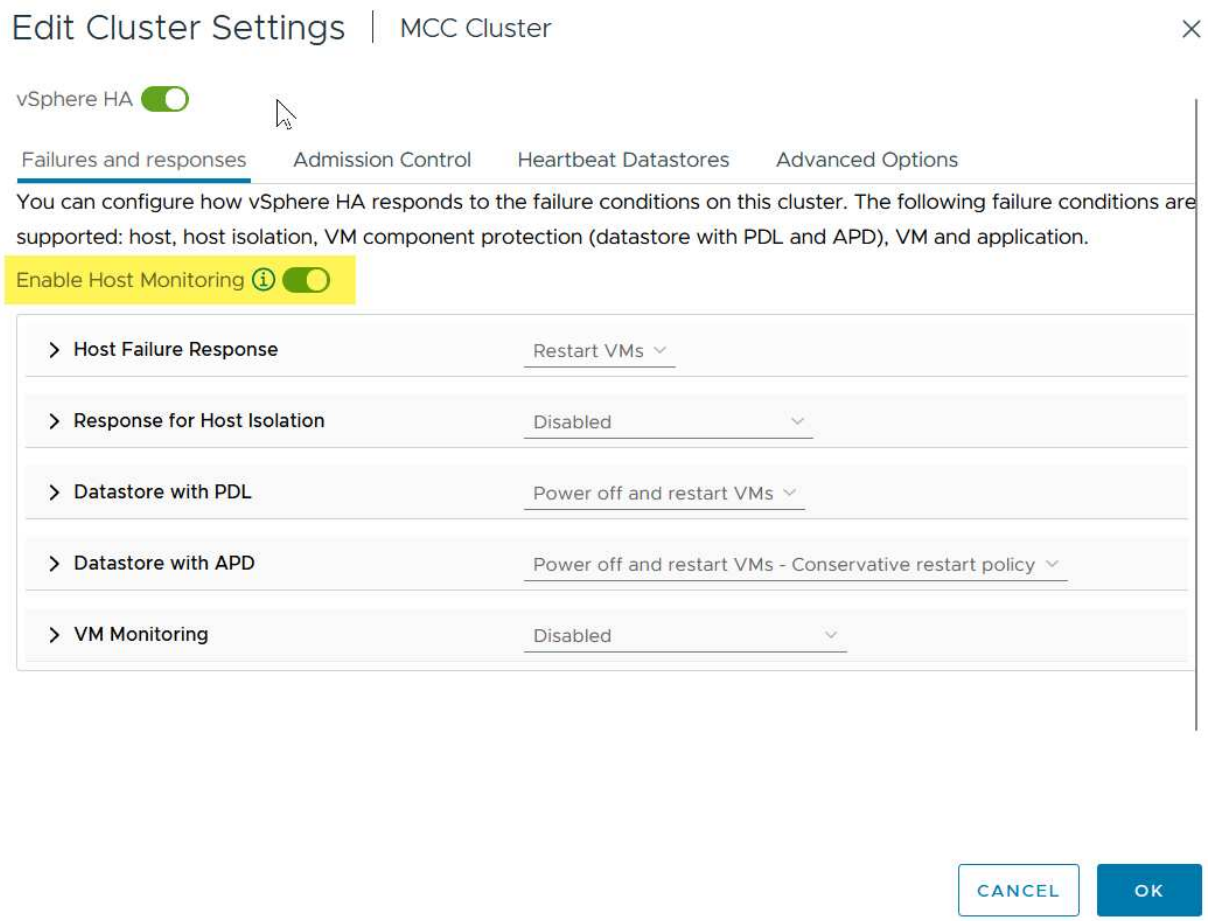
Nota: nulla di quanto contenuto nel presente documento sostituisce "Procedure consigliate per VMware vSphere Metro Storage Cluster"

Per configurare un cluster ha, completare i seguenti passaggi:

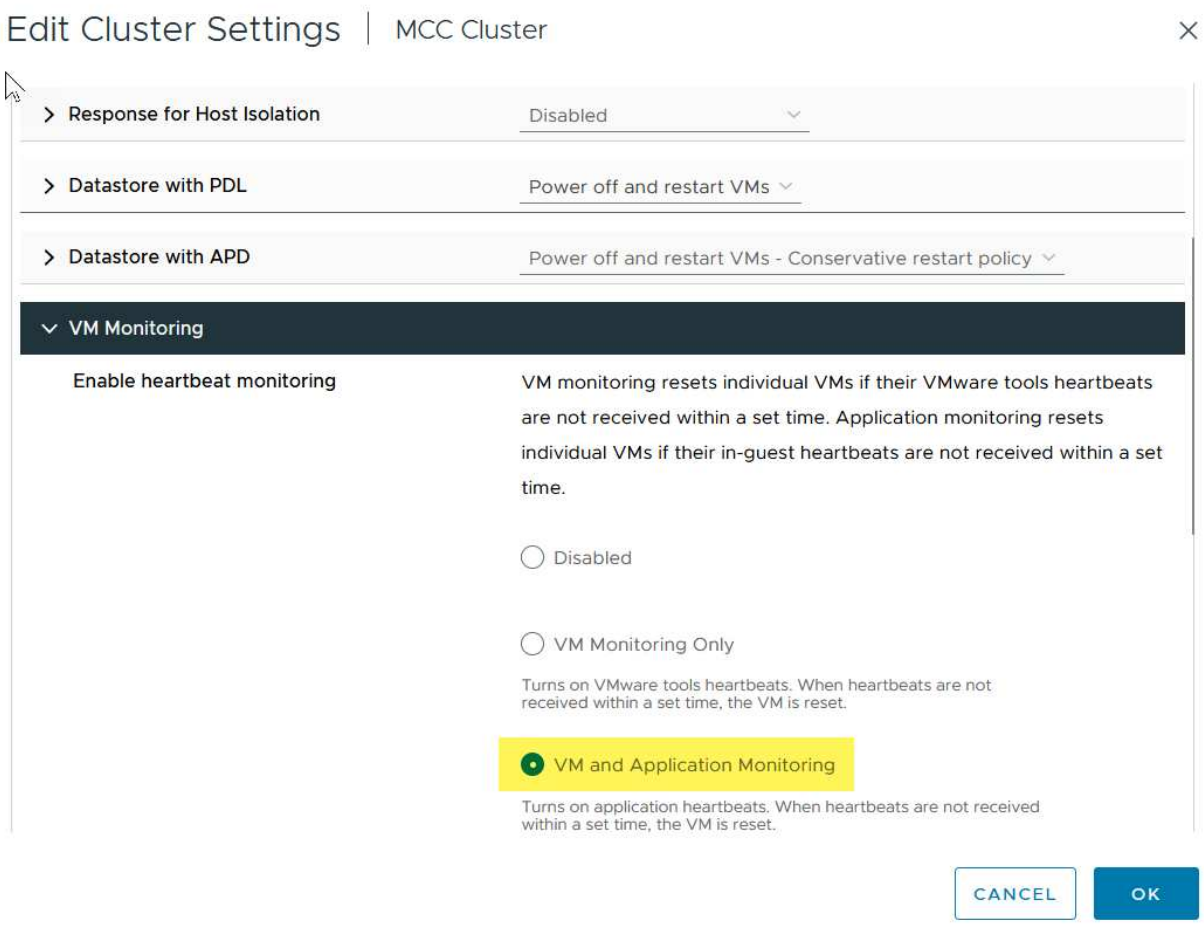
1. Connettersi all'interfaccia utente di vCenter.
2. In host e cluster, individuare il data center in cui si desidera creare il cluster ha.
3. Fare clic con il pulsante destro del mouse sull'oggetto del data center e selezionare nuovo cluster. In base alle nozioni di base, assicurarsi di aver abilitato vSphere DRS e vSphere ha. Completare la procedura guidata.



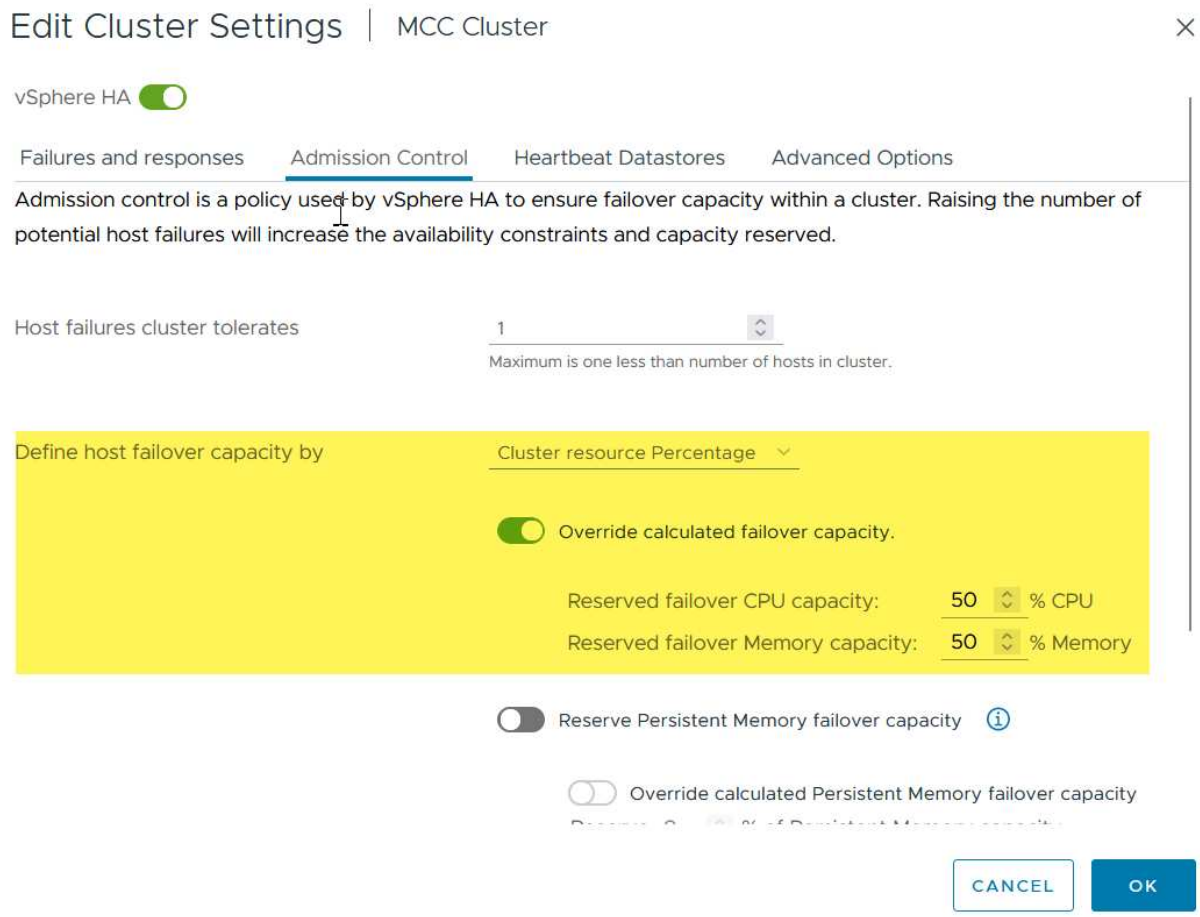
1. Selezionare il cluster e accedere alla scheda di configurazione. Selezionare vSphere ha e fare clic su Modifica.
2. In monitoraggio host, selezionare l'opzione attiva monitoraggio host.



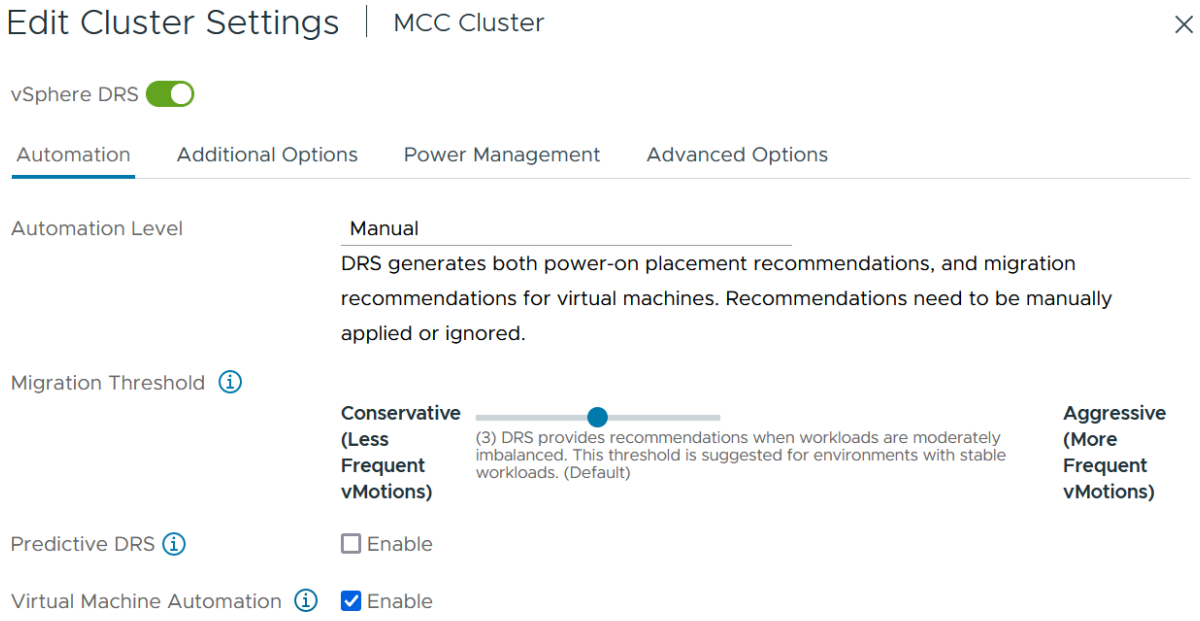
1. Nella scheda guasti e risposte, in monitoraggio VM, selezionare l'opzione solo monitoraggio VM o monitoraggio VM e applicazione.



1. In controllo ammissione, impostare l'opzione di controllo ammissione ha su Cluster Resource Reserve; utilizzare 50% CPU/MEM.



1. Fare clic su "OK".
2. Selezionare DRS e fare clic su MODIFICA.
3. Impostare il livello di automazione su manuale, a meno che non sia richiesto dalle applicazioni.



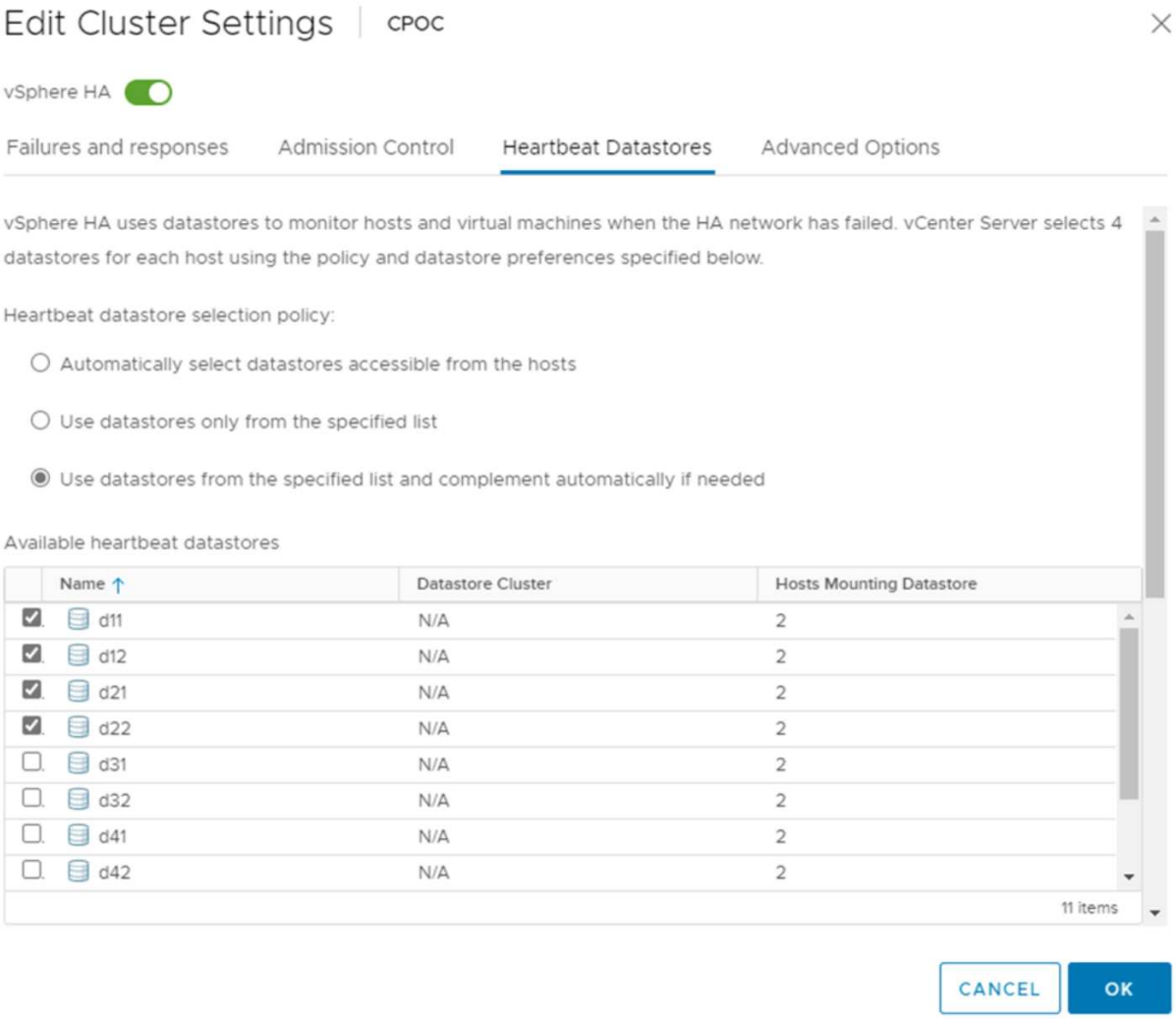
1. Abilitare la protezione dei componenti VM, fare riferimento a. "docs.vmware.com".
2. Le seguenti impostazioni aggiuntive di vSphere ha sono consigliate per vMSC con MCC:

Guasto	Risposta
Errore host	Riavviare le VM
Isolamento degli host	Disattivato
Datastore con perdita permanente di dispositivi (PDL)	Spegnere e riavviare le macchine virtuali
Datastore con tutti i percorsi verso il basso (APD)	Spegnere e riavviare le macchine virtuali
L'ospite non batte il cuore	Ripristinare le VM
Policy di riavvio della VM	Determinato dall'importanza della VM
Risposta per l'isolamento dell'host	Arrestare e riavviare le VM
Risposta per il datastore con PDL	Spegnere e riavviare le macchine virtuali
Risposta per datastore con APD	Spegnere e riavviare le macchine virtuali (conservative)
Ritardo del failover delle macchine virtuali per APD	3 minuti
Risposta per il ripristino APD con timeout APD	Disattivato
Sensibilità di monitoraggio VM	Preimpostazione alta

Configurare gli archivi dati per Heartbeating

VSphere ha utilizza i datastore per monitorare gli host e le macchine virtuali in caso di guasto alla rete di gestione. È possibile configurare in che modo vCenter seleziona i datastore heartbeat. Per configurare gli archivi dati per il heartbeat, completare i seguenti passaggi:

1. Nella sezione Heartbeating del datastore, selezionare Usa archivi dati dall'elenco specificato e completare automaticamente se necessario.
2. Seleziona i datastore che desideri utilizzare vCenter da entrambi i siti e premi OK.

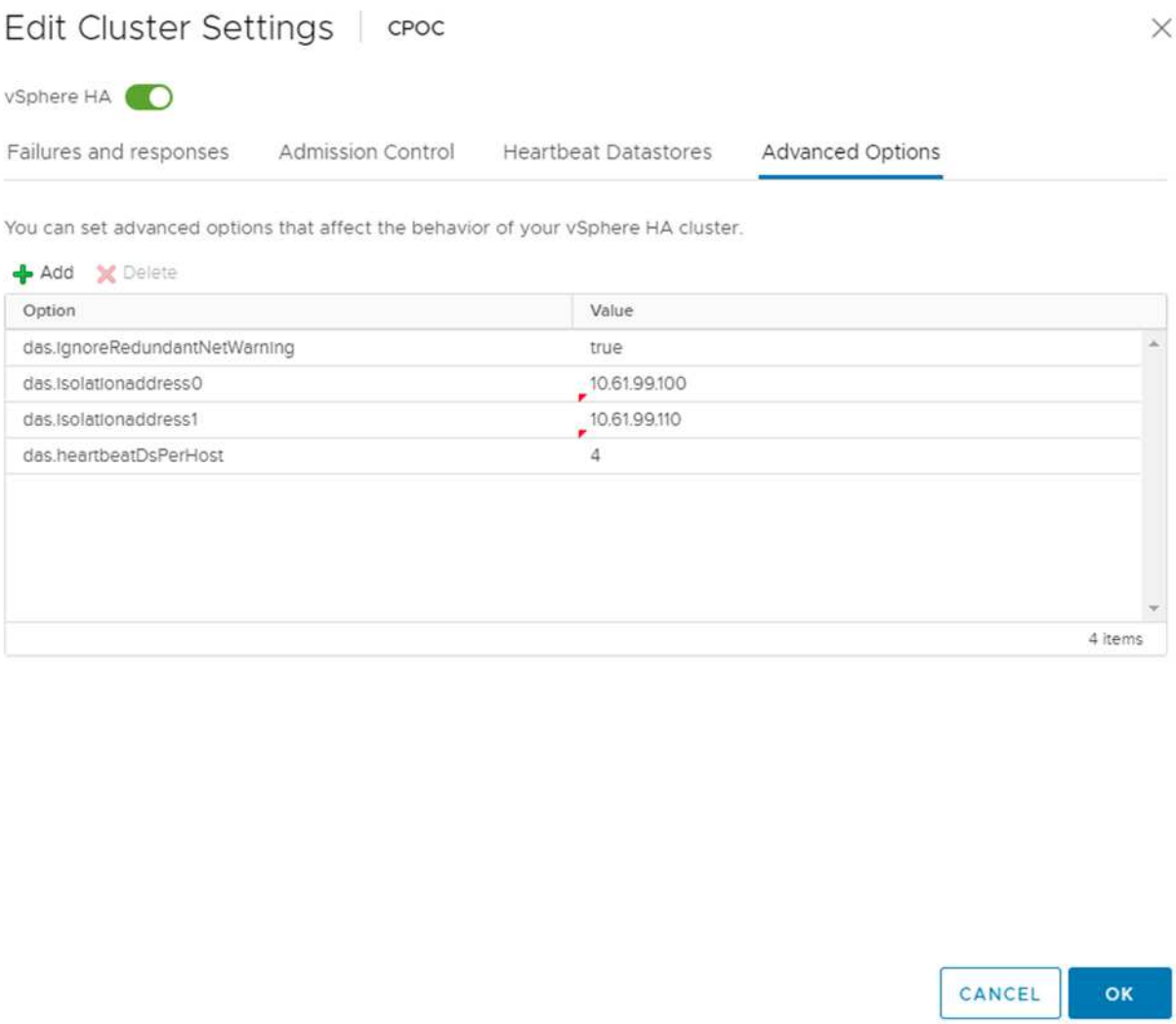


Configurare le opzioni avanzate

Rilevamento errori host

Gli eventi di isolamento si verificano quando gli host all'interno di un cluster ha perdono la connettività alla rete o ad altri host nel cluster. Per impostazione predefinita, vSphere ha utilizzerà il gateway predefinito per la propria rete di gestione come indirizzo di isolamento predefinito. Tuttavia, è possibile specificare indirizzi di isolamento aggiuntivi per l'host al ping per determinare se deve essere attivata una risposta di isolamento. Aggiungere due IP di isolamento in grado di eseguire il ping, uno per sito. Non utilizzare l'indirizzo IP del gateway. L'impostazione avanzata vSphere ha utilizzata è das.isolationaddress. A tale scopo, è possibile utilizzare gli indirizzi IP ONTAP o Mediator.

Fare riferimento a. "core.vmware.com" per ulteriori informazioni.



L'aggiunta di un'impostazione avanzata denominata das.heartbeatDsPerHost può aumentare il numero di datastore heartbeat. Utilizzare quattro datastore heartbeat (HB DSS), due per sito. Utilizzare l'opzione "Select from List but complent" (Seleziona da elenco ma complimento). Questo è necessario perché se un sito non funziona, è necessario ancora due HB DSS. Tuttavia, questi elementi non devono essere protetti con la sincronizzazione attiva di MCC o SnapMirror.

Fare riferimento a. "core.vmware.com" per ulteriori informazioni.

Affinità con VMware DRS per NetApp MetroCluster

In questa sezione vengono creati gruppi DRS per VM e host per ciascun sito/cluster nell'ambiente MetroCluster. Quindi configuriamo le regole VM\host per allineare l'affinità dell'host VM con le risorse di storage locali. Ad esempio, il sito A fa parte del gruppo VM sitea_VM e gli host del sito A appartengono al gruppo host sitea_hosts. Successivamente, in VM\host Rules, si afferma che sitea_vm deve essere eseguito sugli host in sitea_hosts.

Best practice

- NetApp consiglia vivamente la specifica **deve essere eseguita sugli host nel gruppo** piuttosto che sulla specifica **deve essere eseguita sugli host nel gruppo**. In caso di guasto dell'host del sito A, è necessario riavviare le macchine virtuali del sito A sugli host del sito B attraverso vSphere ha, ma quest'ultima specifica non consente all'ha di riavviare le macchine virtuali sul sito B perché è una regola rigida. La

specifica precedente è una regola debole e viene violata in caso di ha, abilitando in tal modo la disponibilità anziché le prestazioni.

Nota: è possibile creare un allarme basato su eventi che viene attivato quando una macchina virtuale viola una regola di affinità VM-host. Nel client vSphere, aggiungere un nuovo allarme per la macchina virtuale e selezionare "VM viola la regola di affinità VM-host" come trigger dell'evento. Per ulteriori informazioni sulla creazione e la modifica degli allarmi, fare riferimento a. "Monitoraggio e performance di vSphere" documentazione.

Creare gruppi host DRS

Per creare gruppi di host DRS specifici per il sito A e il sito B, attenersi alla seguente procedura:

1. Nel client web vSphere, fare clic con il pulsante destro del mouse sul cluster nell'inventario e selezionare Impostazioni.
2. Fare clic su VM\host Groups.
3. Fare clic su Aggiungi.
4. Digitare il nome del gruppo (ad esempio, sitea_hosts).
5. Dal menu tipo, selezionare Gruppo host.
6. Fare clic su Aggiungi e selezionare gli host desiderati dal sito A, quindi fare clic su OK.
7. Ripetere questi passaggi per aggiungere un altro gruppo di host per il sito B.
8. Fare clic su OK.

Creare gruppi DRS VM

Per creare gruppi di macchine virtuali DRS specifici per il sito A e il sito B, attenersi alla seguente procedura:

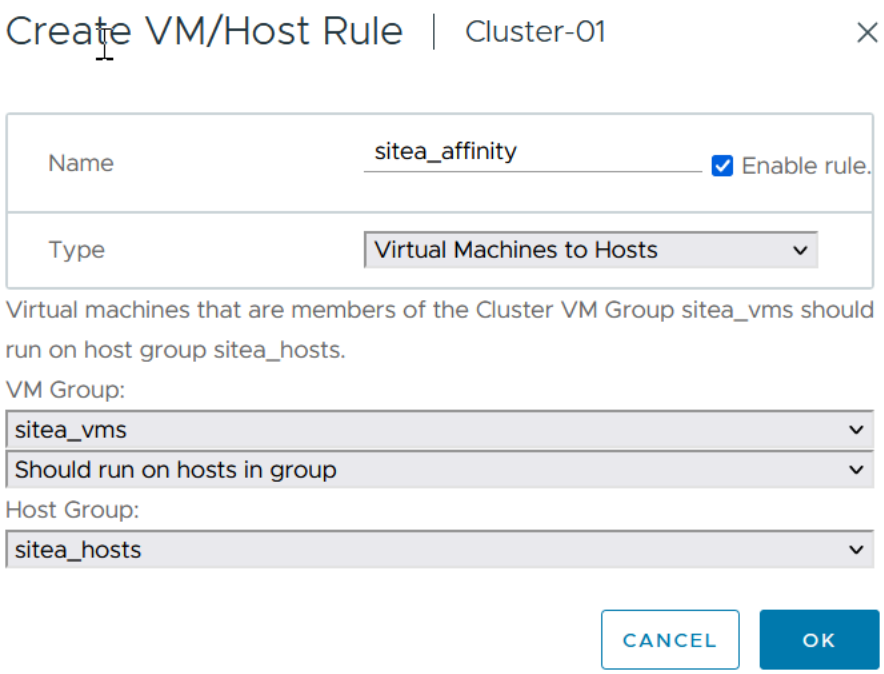
1. Nel client web vSphere, fare clic con il pulsante destro del mouse sul cluster nell'inventario e selezionare Impostazioni.
2. Fare clic su VM\host Groups.
3. Fare clic su Aggiungi.
4. Digitare il nome del gruppo (ad esempio, sitea_vm).
5. Dal menu tipo, selezionare Gruppo VM.
6. Fare clic su Add (Aggiungi) e selezionare le VM desiderate dal sito A, quindi fare clic su OK.
7. Ripetere questi passaggi per aggiungere un altro gruppo di host per il sito B.
8. Fare clic su OK.

Crea regole host VM

Per creare regole di affinità DRS specifiche per il sito A e il sito B, completare i seguenti passaggi:

1. Nel client web vSphere, fare clic con il pulsante destro del mouse sul cluster nell'inventario e selezionare Impostazioni.
2. Fare clic su VM\host Rules.
3. Fare clic su Aggiungi.
4. Digitare il nome della regola (ad esempio, sitea_Affinity).

5. Verificare che l'opzione Enable Rule (attiva regola) sia selezionata.
6. Dal menu Type (tipo), selezionare Virtual Machines to hosts (macchine virtuali a host).
7. Selezionare il gruppo VM (ad esempio, sitea_vm).
8. Selezionare il gruppo host (ad esempio, sitea_hosts).
9. Ripetere questi passaggi per aggiungere un'altra VM\regola host per il sito B.
10. Fare clic su OK.



VMware vSphere Storage DRS per NetApp MetroCluster

Creare cluster di datastore

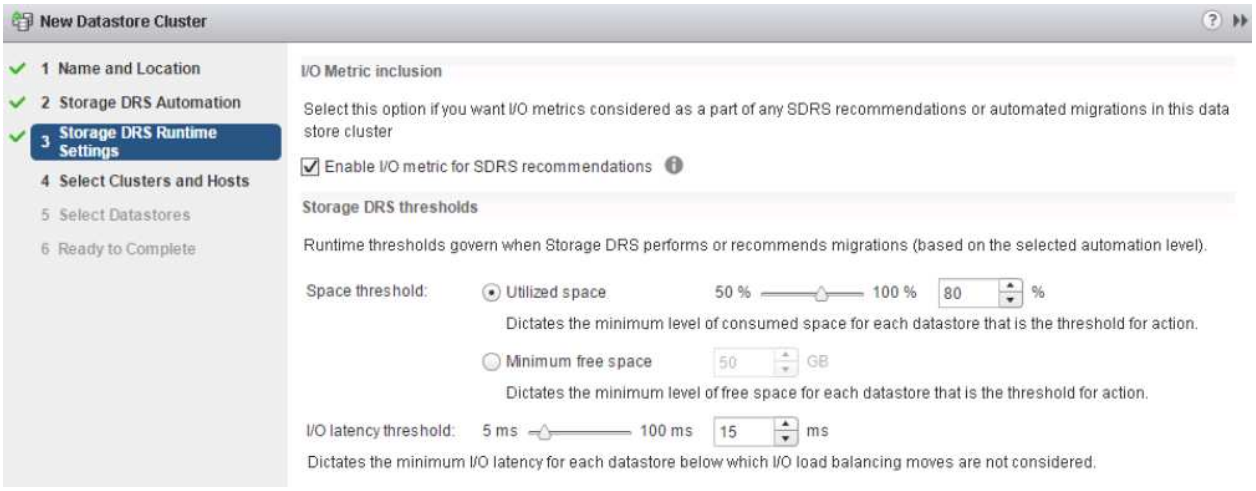
Per configurare un cluster di datastore per ciascun sito, attenersi alla seguente procedura:

1. Utilizzando il client web vSphere, individuare il data center in cui risiede il cluster ha in Storage.
2. Fare clic con il pulsante destro del mouse sull'oggetto del data center e selezionare Storage > New Datastore Cluster.
3. Selezionare l'opzione Turn ON Storage DRS (ATTIVA DRS archiviazione) e fare clic su Next (Avanti).
4. Impostare tutte le opzioni su Nessuna automazione (modalità manuale) e fare clic su Avanti.

Best practice

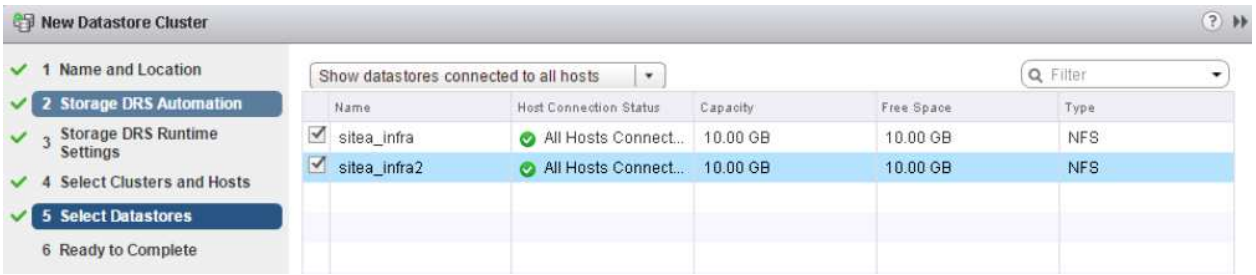
- NetApp consiglia di configurare i DRS dello storage in modalità manuale, in modo che l'amministratore possa decidere e controllare quando è necessario eseguire le migrazioni.

1. Verificare che la casella di controllo Enable i/o Metric for SDRS Recommendations (Abilita metriche i/o per raccomandazioni SDRS) sia selezionata; le impostazioni metriche possono essere lasciate con i valori predefiniti.



1. Selezionare il cluster ha e fare clic su Next.

1. Selezionare gli archivi dati appartenenti al sito A e fare clic su Avanti.



1. Rivedere le opzioni e fare clic su fine.
2. Ripetere questa procedura per creare il cluster di datastore del sito B e verificare che siano selezionati solo i datastore del sito B.

Disponibilità di vCenter Server

Le appliance vCenter Server (VCSA) devono essere protette con vCenter ha. VCenter ha ti consente di implementare due VCSA in una coppia ha Active-passive. Uno in ogni dominio di errore. Puoi leggere ulteriori informazioni su vCenter ha all'indirizzo "docs.vmware.com".

Resilienza per eventi pianificati e non pianificati

NetApp MetroCluster e SnapMirror Active Sync sono potenti strumenti che migliorano l'alta disponibilità e le operazioni senza interruzioni dell'hardware NetApp e del software ONTAP®.

Questi strumenti garantiscono una protezione a livello di sito per l'intero ambiente di storage, garantendo che i tuoi dati siano sempre disponibili. Che si stiano utilizzando server standalone, cluster di server ad alta disponibilità, container Docker o server virtualizzati, la tecnologia NetApp permette di conservare perfettamente la disponibilità dello storage in caso di black-out totale causato da black-out, raffreddamento o connettività di rete, arresto dello storage array o errori operativi.

La sincronizzazione attiva di MetroCluster e SnapMirror offre tre metodi di base per la continuità dei dati in caso di eventi pianificati o non pianificati:

- Componenti ridondanti per la protezione contro i guasti a un singolo componente
- Takeover locale di ha in caso di eventi che colpiscono un singolo controller
- Protezione completa del sito: Rapida ripresa del servizio mediante il trasferimento dello storage e dell'accesso client dal cluster di origine al cluster di destinazione

Ciò significa che le operazioni continuano senza problemi in caso di guasto a un singolo componente e vengono ripristinate automaticamente al funzionamento ridondante una volta sostituito il componente guasto.

Tutti i cluster ONTAP, ad eccezione dei cluster a nodo singolo (in genere versioni software-defined, come ad esempio ONTAP Select), offrono funzionalità di ha integrate chiamate takeover e giveback. Ciascun controller del cluster è accoppiato con un altro controller in modo da formare una coppia ha. Queste coppie garantiscono che ogni nodo sia connesso localmente allo storage.

Il takeover è un processo automatizzato in cui un nodo assume il controllo dello storage dell'altro per la gestione dei servizi dati. Giveback è il processo inverso che ripristina il normale funzionamento. Il takeover può essere pianificato, ad esempio durante la manutenzione hardware o gli upgrade della ONTAP, o non pianificato, derivante da un nodo di panico o da un guasto dell'hardware.

Durante un takeover, le interfacce logiche NAS (Network Attached Storage) nelle configurazioni MetroCluster eseguono automaticamente il failover. Tuttavia, le LIF (SAN) di Storage Area Network non subiscono failover e continueranno a utilizzare il percorso diretto dei LUN (Logical Unit Number).

Per ulteriori informazioni sul takeover e lo sconto ha, consulta la "Panoramica sulla gestione delle coppie HA". È importante notare che questa funzionalità non è specifica per MetroCluster o SnapMirror Active Sync.

Lo switchover del sito con MetroCluster viene eseguito quando un sito è offline o come attività pianificata per la manutenzione di un intero sito. Il sito rimanente presuppone la proprietà delle risorse storage (dischi e aggregati) del cluster offline e le SVM del sito guasto vengono messe online e riavviate nel sito di disaster recovery, preservando la loro identità completa per l'accesso client e host.

Con la sincronizzazione attiva di SnapMirror, poiché entrambe le copie vengono utilizzate contemporaneamente in modo attivo, gli host esistenti continueranno a funzionare. Il NetApp Mediator è

necessario per garantire che il failover del sito avvenga correttamente.

Scenari di errore per vMSC con MCC

Nelle sezioni seguenti vengono illustrati i risultati attesi da vari scenari di guasto con i sistemi vMSC e NetApp MetroCluster.

Errore singolo percorso di storage

In questo scenario, se componenti come la porta HBA, la porta di rete, la porta dello switch dati front-end o un cavo FC o Ethernet si guastano, quel particolare percorso al dispositivo di storage viene contrassegnato come inattivo dall'host ESXi. Se vengono configurati diversi percorsi per il dispositivo storage fornendo resilienza alla porta HBA/rete/switch, ESXi esegue uno switchover del percorso. Durante questo periodo, le macchine virtuali rimangono in esecuzione senza alcun impatto, perché la disponibilità dello storage viene garantita attraverso l'offerta di più percorsi al dispositivo di storage.

Nota: in questo scenario non vi è alcun cambiamento nel comportamento di MetroCluster, e tutti i datastore continuano ad essere intatti dai rispettivi siti.

Best practice

Negli ambienti in cui vengono utilizzati volumi NFS/iSCSI, NetApp consiglia di avere almeno due uplink di rete configurati per la porta vmkernel NFS nel vSwitch standard e lo stesso nel gruppo di porte in cui è mappata l'interfaccia vmkernel NFS per il vSwitch distribuito. Il raggruppamento NIC può essere configurato in modalità Active-Active o Active-standby.

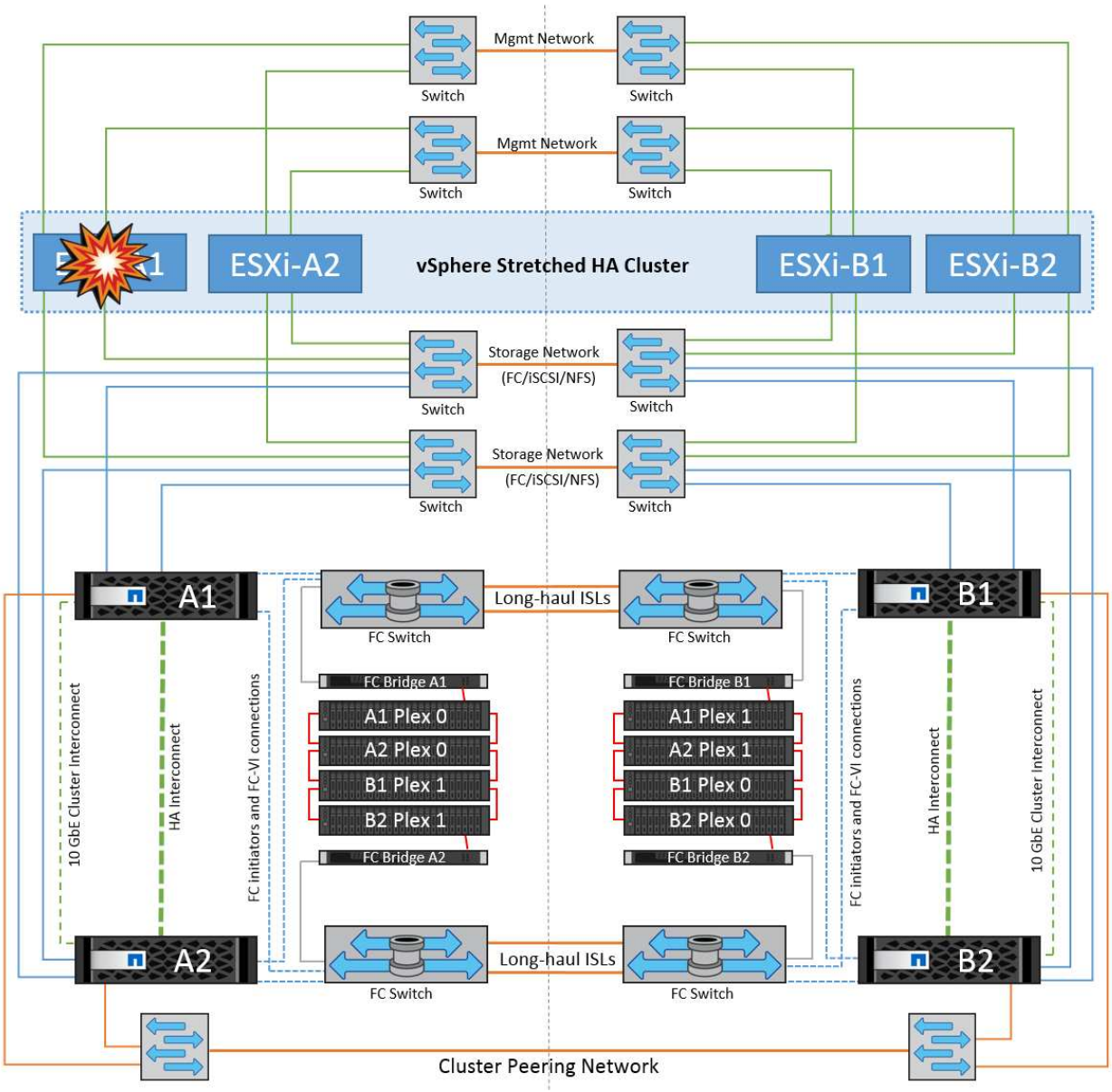
Inoltre, per i LUN iSCSI, il multipathing deve essere configurato legando le interfacce vmkernel agli adattatori di rete iSCSI. Per ulteriori informazioni, fai riferimento alla documentazione dello storage vSphere.

Best practice

Negli ambienti in cui vengono utilizzate le LUN Fibre Channel, NetApp consiglia di disporre di almeno due HBA, che garantiscono resilienza a livello di HBA/porta. NetApp consiglia inoltre di utilizzare lo zoning a destinazione singola come Best practice per la configurazione dello zoning.

È necessario utilizzare Virtual Storage Console (VSC) per impostare policy di multipathing, perché imposta policy per tutti i dispositivi storage NetApp nuovi ed esistenti.

Errore host ESXi singolo



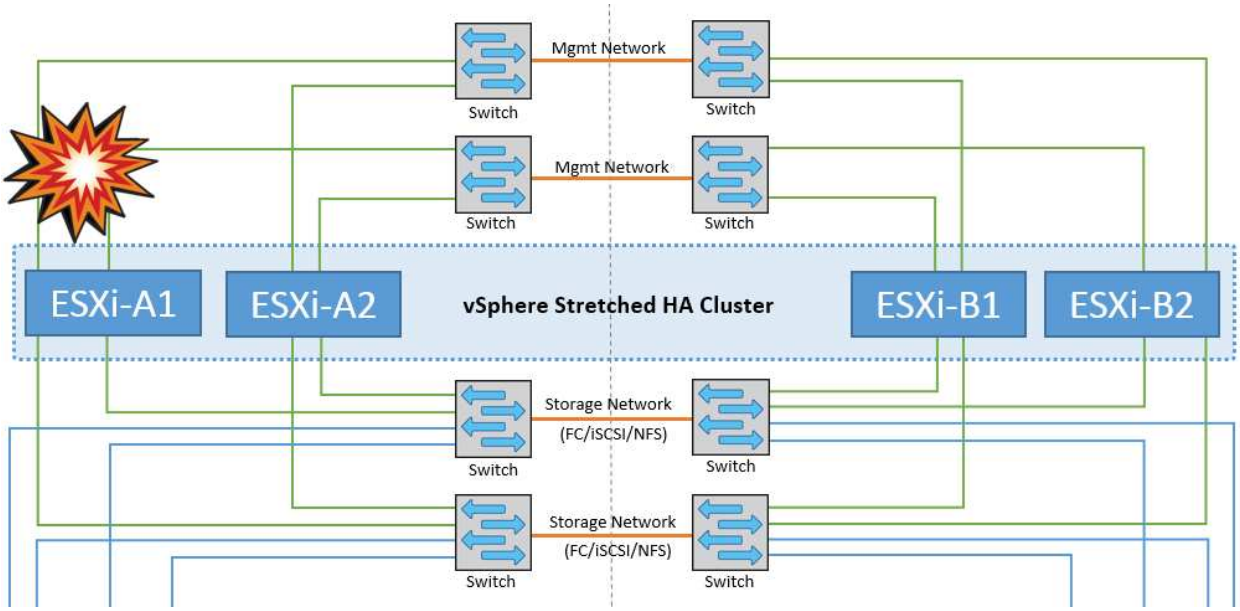
In questo scenario, se si verifica un guasto dell'host ESXi, il nodo master nel cluster VMware ha rileva il guasto dell'host in quanto non riceve più gli heartbeat di rete. Per determinare se l'host è effettivamente inattivo o solo una partizione di rete, il nodo master monitora gli heartbeat del datastore e, se sono assenti, esegue un controllo finale eseguendo il ping degli indirizzi IP di gestione dell'host guasto. Se tutti questi controlli sono negativi, il nodo master dichiara l'host un host guasto e tutte le macchine virtuali in esecuzione su questo host guasto vengono riavviate sull'host rimasto nel cluster.

Se sono state configurate le regole di affinità per DRS VM e host (le VM nel gruppo VM sitea_VM devono eseguire gli host nel gruppo host sitea_hosts), il master ha controlla prima le risorse disponibili nel sito A. Se non ci sono host disponibili nel sito A, il master tenta di riavviare le VM sugli host nel sito B.

È possibile che le macchine virtuali vengano avviate sugli host ESXi nell'altro sito se è presente un vincolo di risorse nel sito locale. Tuttavia, le regole di affinità definite per DRS VM e host verranno corrette in caso di violazione di regole mediante la migrazione delle macchine virtuali a qualsiasi host ESXi rimasto nel sito locale. Nei casi in cui DRS è impostato su manuale, NetApp consiglia di richiamare DRS e applicare le raccomandazioni per correggere il posizionamento della macchina virtuale.

In questo scenario, non vi sono cambiamenti nel comportamento di MetroCluster e tutti i datastore continuano a essere intatti dai rispettivi siti.

Isolamento dell'host ESXi

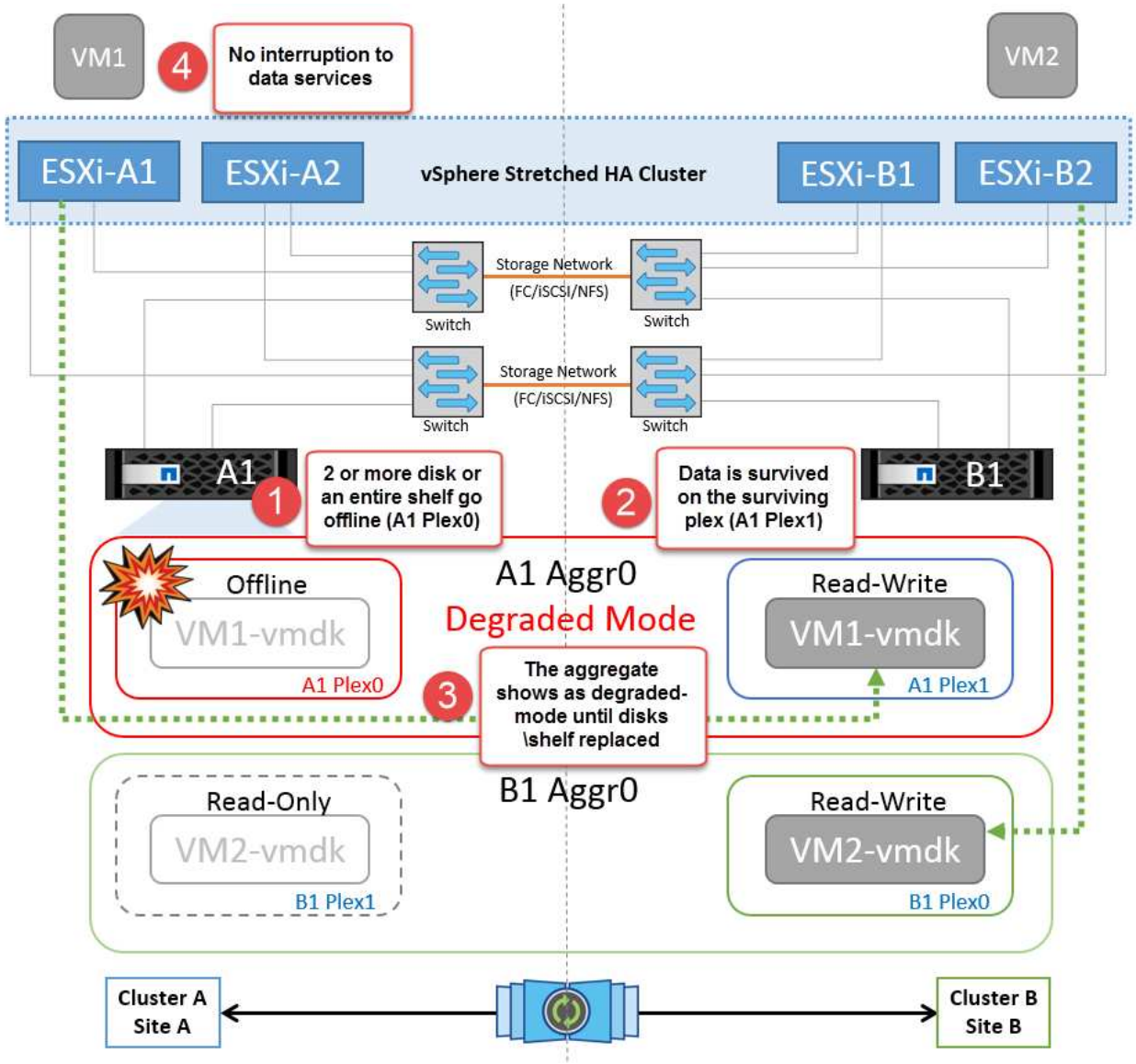


In questo scenario, se la rete di gestione dell'host ESXi non è attiva, il nodo master nel cluster ha non riceverà alcun heartbeat, pertanto l'host viene isolato nella rete. Per determinare se si è verificato un errore o se è solo isolato, il nodo master inizia a monitorare l'heartbeat del datastore. Se è presente, l'host viene dichiarato isolato dal nodo master. A seconda della risposta di isolamento configurata, l'host può scegliere di spegnere, spegnere le macchine virtuali o persino lasciare accese le macchine virtuali. L'intervallo predefinito per la risposta di isolamento è di 30 secondi.

In questo scenario, non vi sono cambiamenti nel comportamento di MetroCluster e tutti i datastore continuano a essere intatti dai rispettivi siti.

Guasto a shelf di dischi

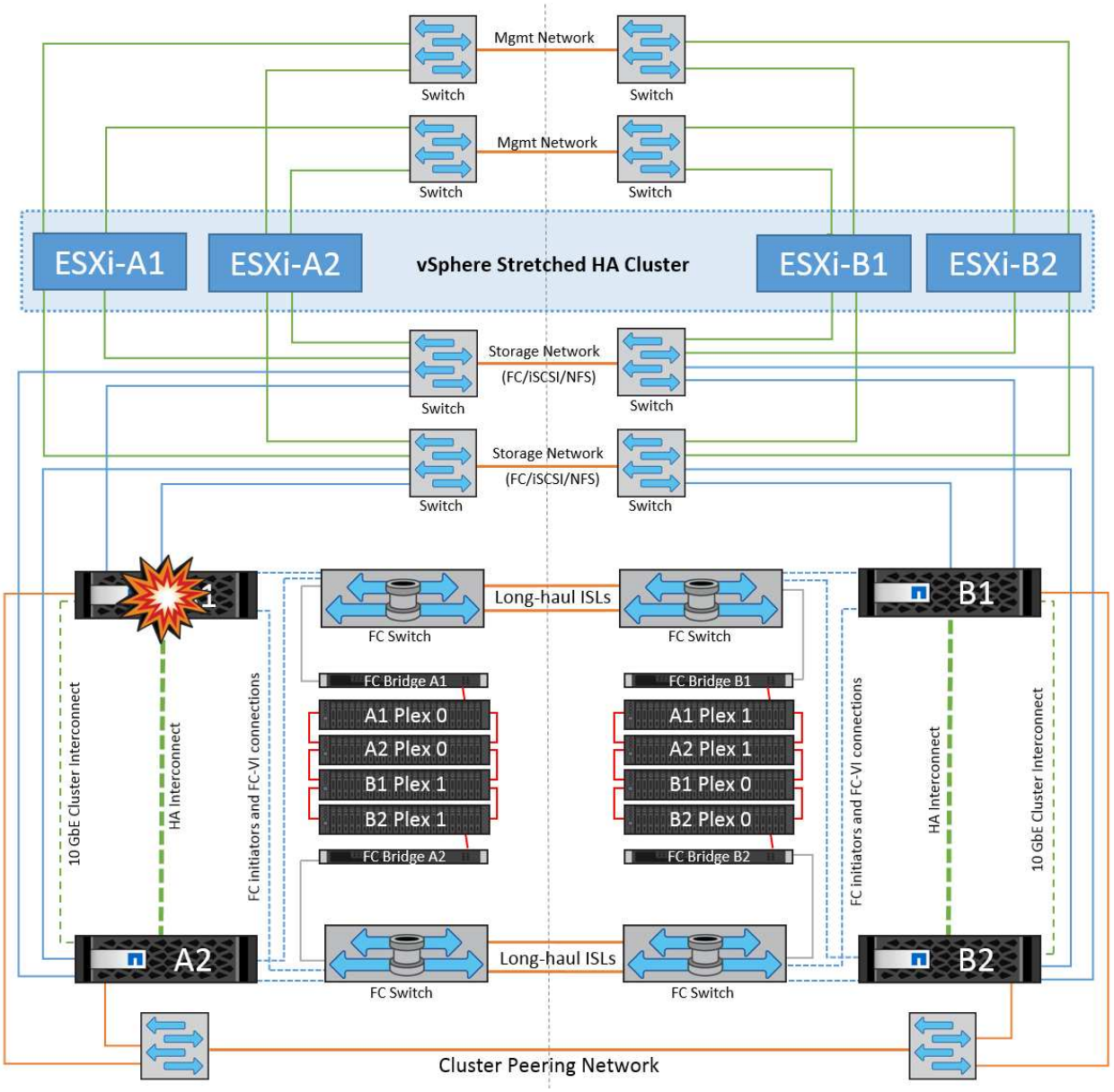
In questo scenario, si verifica un errore di più di due dischi o di un intero shelf. I dati vengono distribuiti dal plesso restante senza alcuna interruzione dei servizi dati. Il guasto del disco potrebbe influire su un plesso locale o remoto. Gli aggregati vengono visualizzati come modalità degradata perché è attivo un solo plesso. Una volta sostituiti i dischi guasti, gli aggregati interessati si risincronizzano automaticamente per ricostruire i dati. Dopo la risincronizzazione, gli aggregati tornano automaticamente alla normale modalità con mirroring. Se più di due dischi all'interno di un singolo gruppo RAID si sono guastati, il plex deve essere ricostruito da zero.



Nota: durante questo periodo, non si verifica alcun impatto sulle operazioni i/o della macchina virtuale, ma le prestazioni sono ridotte a causa dell'accesso ai dati dallo shelf di dischi remoto tramite collegamenti ISL.

Guasto a un singolo storage controller

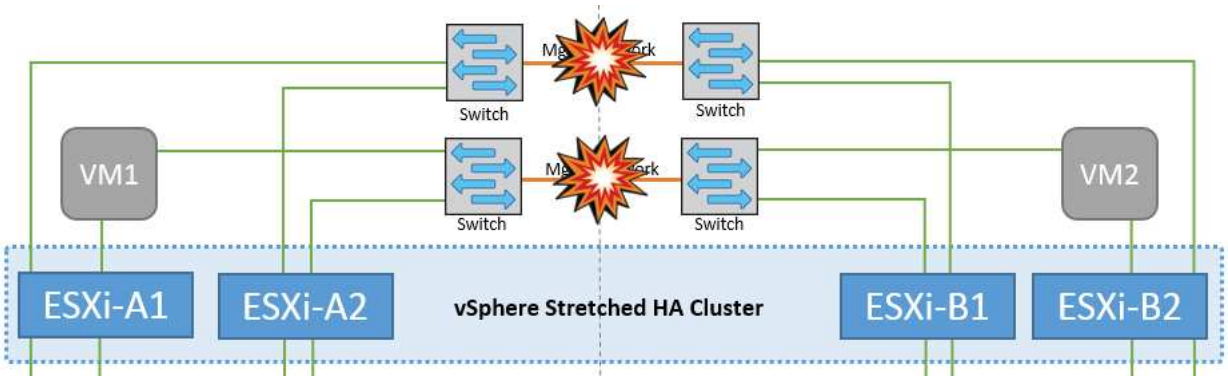
In questo scenario, uno dei due storage controller si guasta in un solo sito. Poiché è presente una coppia ha in ciascun sito, un guasto di un nodo attiva automaticamente il failover sull'altro nodo. Ad esempio, in caso di guasto al nodo A1, il relativo storage e carichi di lavoro vengono trasferiti automaticamente al nodo A2. Le macchine virtuali non saranno interessate perché tutti i plessi rimangono disponibili. I nodi del secondo sito (B1 e B2) non sono interessati. Inoltre, vSphere ha non intraprenderà alcuna azione perché il nodo master nel cluster riceverà comunque gli heartbeat di rete.



Se il failover fa parte di un rolling disaster (il nodo A1 esegue il failover su A2) e si verifica un successivo guasto di A2 o il guasto completo del sito A, è possibile eseguire lo switchover in seguito a un disastro nel sito B.

Errori del collegamento interswitch

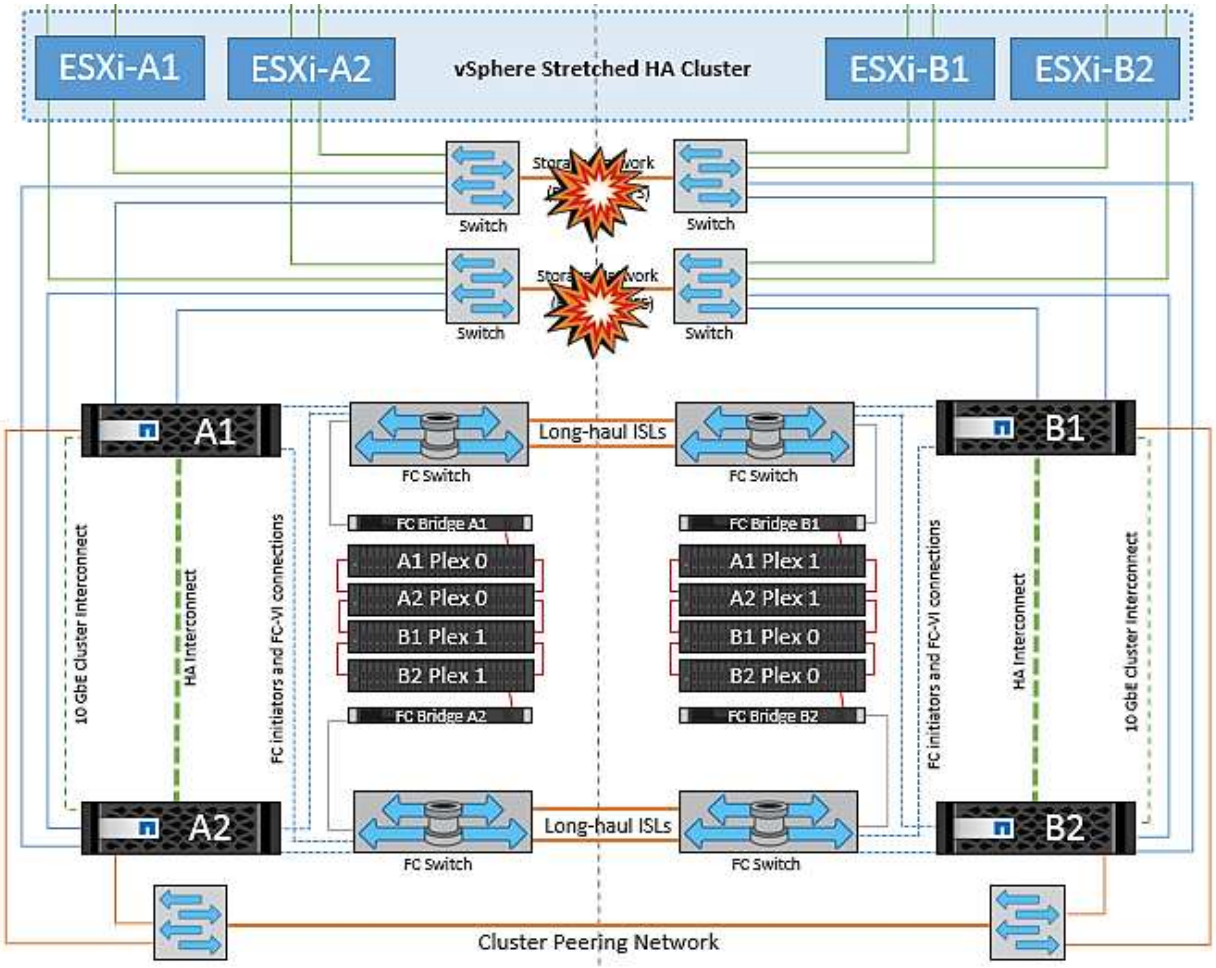
Errore collegamento interswitch sulla rete di gestione



In questo scenario, se i collegamenti ISL nella rete di gestione host front-end si guastano, gli host ESXi nel sito A non saranno in grado di comunicare con gli host ESXi nel sito B. Ciò determina una partizione di rete poiché gli host ESXi in un determinato sito non sono in grado di inviare gli heartbeat di rete al nodo master nel cluster ha. Come tale, ci saranno due segmenti di rete a causa della partizione e vi sarà un nodo master in ogni segmento che proteggerà le VM da guasti host all'interno del sito specifico.

Nota: durante questo periodo, le macchine virtuali rimangono in esecuzione e non vi è alcuna modifica nel comportamento di MetroCluster in questo scenario. Tutti i datastore continuano a essere intatti dai rispettivi siti.

Errore collegamento interswitch sulla rete di storage

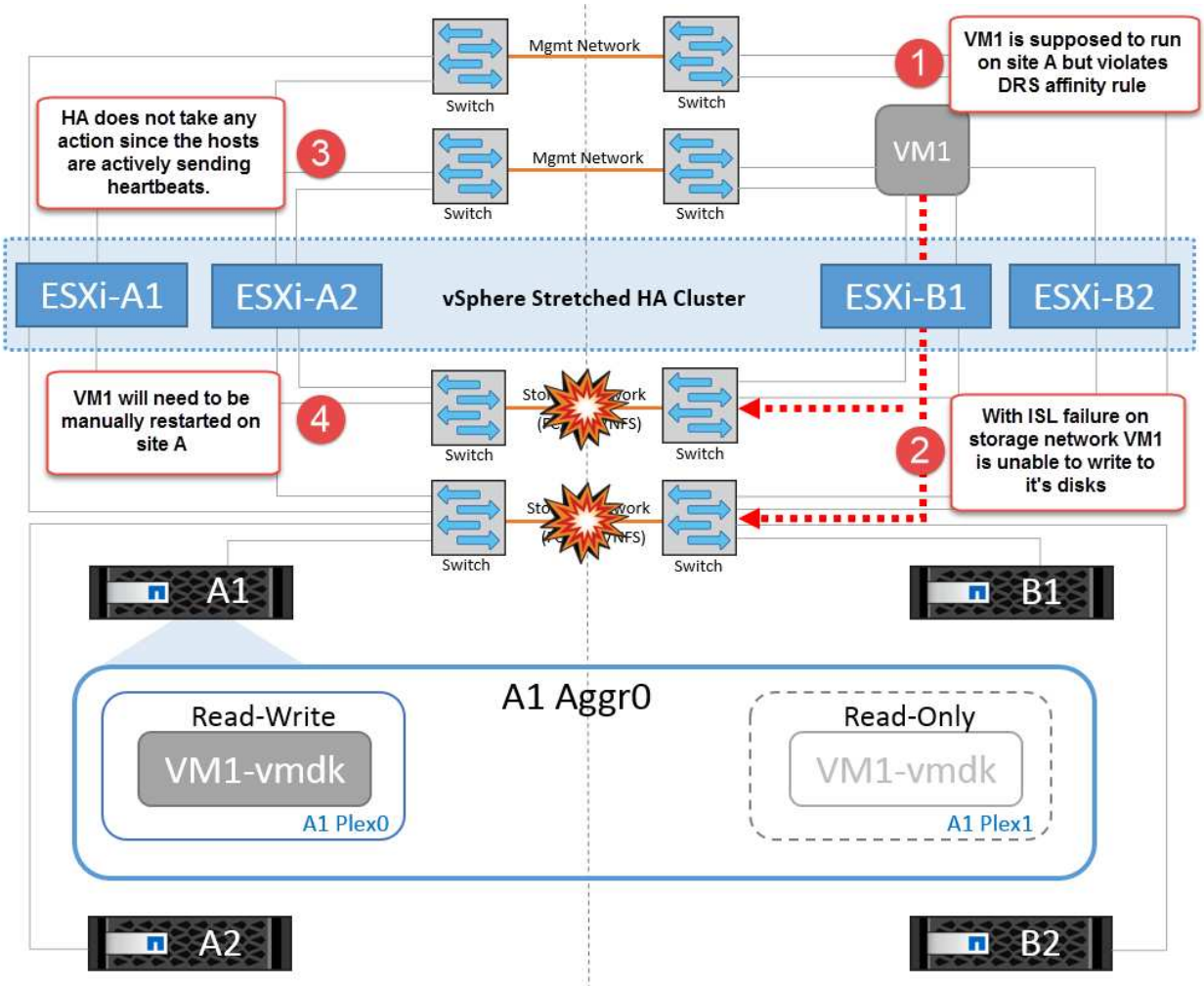


In questo scenario, se si verifica un errore nei collegamenti ISL nella rete di storage backend, gli host sul sito A perderanno l'accesso ai volumi di storage o alle LUN del cluster B nel sito B e viceversa. Le regole VMware DRS sono definite in modo che l'affinità tra il sito host e il sito di storage faciliti l'esecuzione delle macchine

virtuali senza impatti all'interno del sito.

Durante questo periodo, le macchine virtuali rimangono in esecuzione nei rispettivi siti e in questo scenario non si verifica alcuna modifica nel comportamento di MetroCluster. Tutti i datastore continuano a essere intatti dai rispettivi siti.

Se per qualche motivo è stata violata la regola di affinità (ad esempio VM1, che doveva essere eseguito dal sito A in cui i dischi risiedono sui nodi del cluster locale A vengono eseguiti su un host nel sito B), il disco della macchina virtuale può essere acceduto in remoto tramite i link ISL. A causa di un errore del collegamento ISL, VM1 in esecuzione nel sito B non sarebbe in grado di scrivere sui propri dischi perché i percorsi del volume di storage non sono attivi e quella particolare macchina virtuale non è attiva. In queste situazioni, VMware ha non intraprende alcuna azione poiché gli host stanno inviando heartbeat. Tali macchine virtuali devono essere spente e attivate manualmente nei rispettivi siti. La figura seguente illustra una VM che viola una regola di affinità DRS.

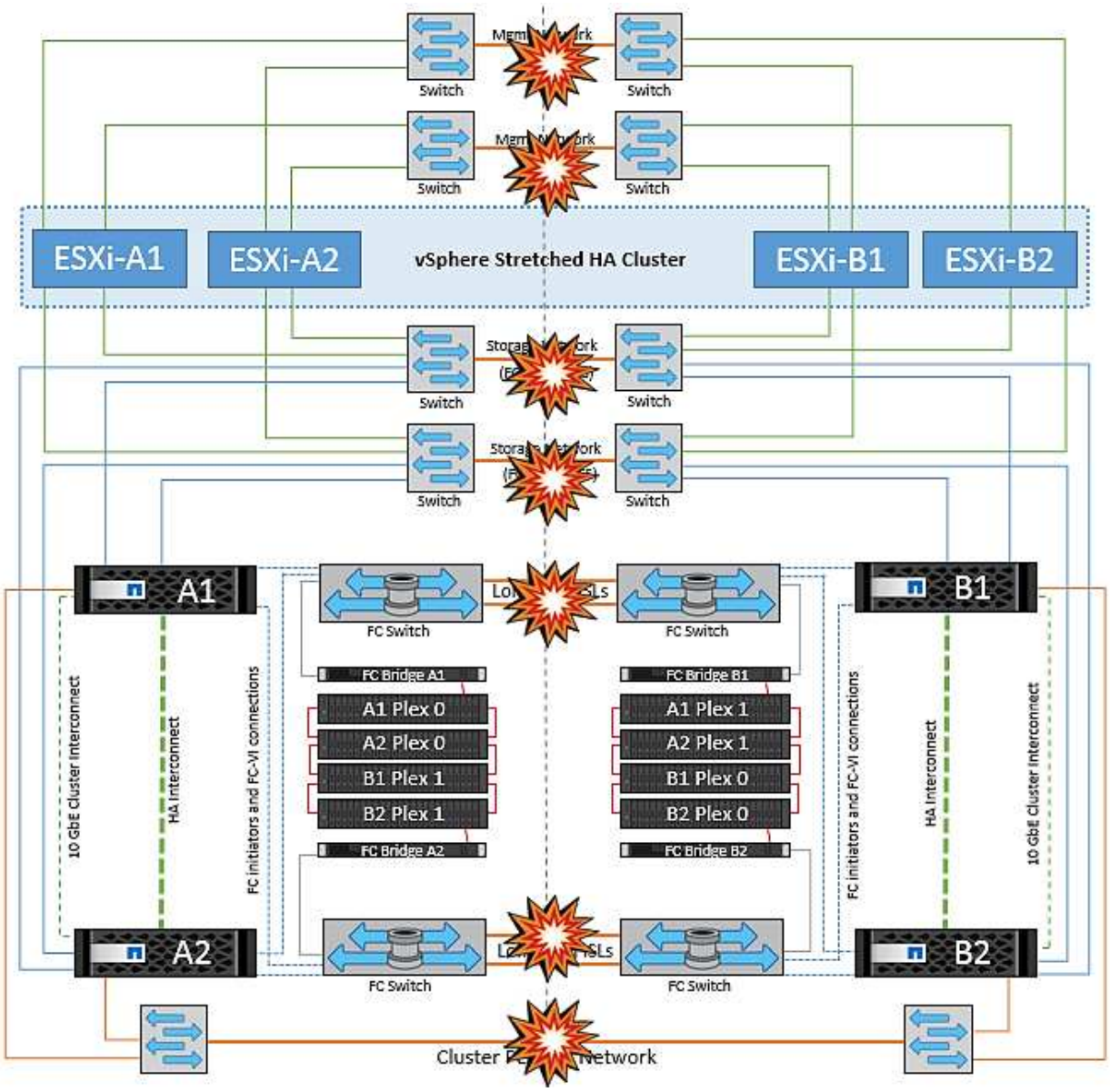


Guasto a tutti gli interswitch o partizione completa del data center

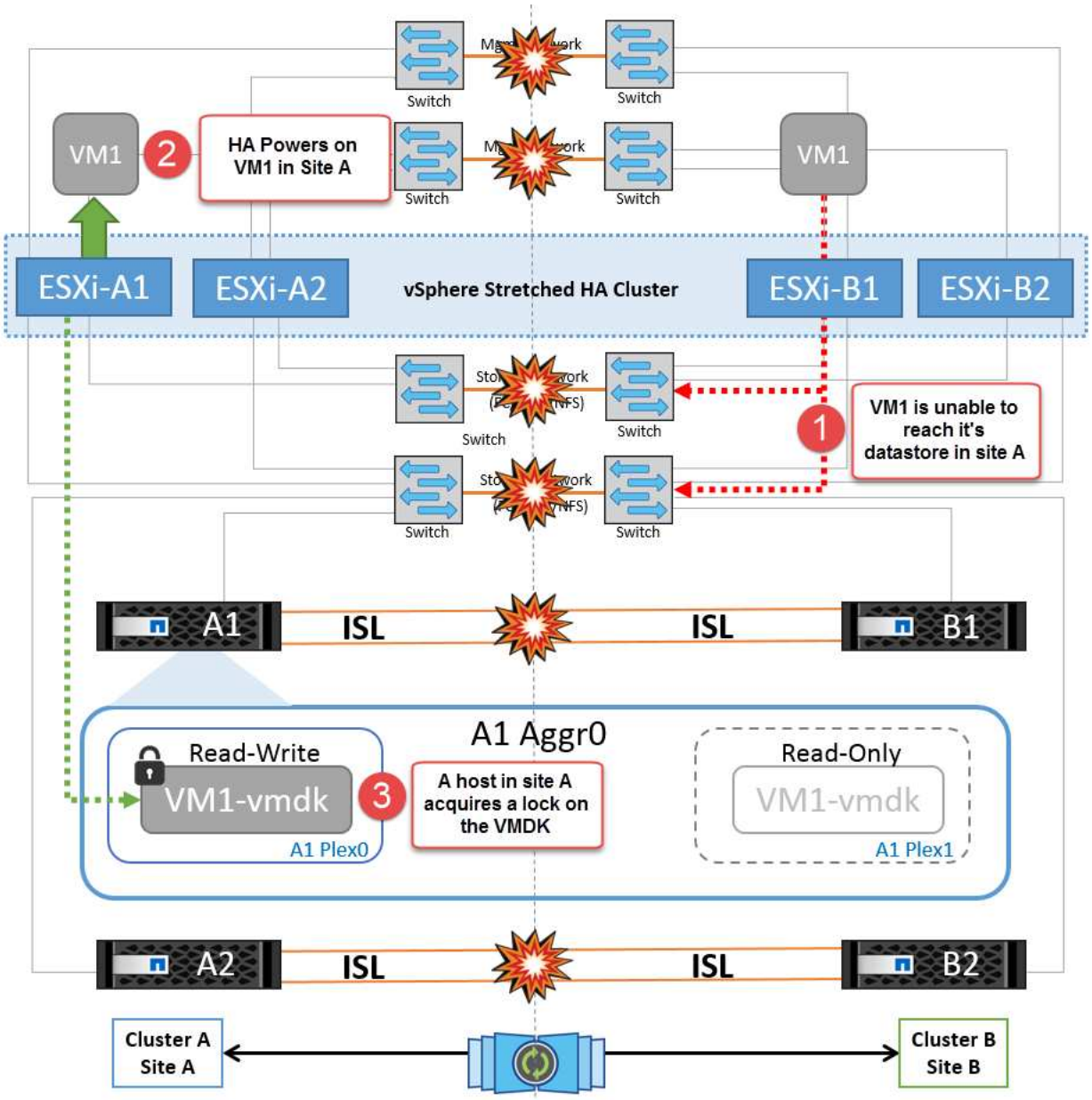
In questo scenario, tutti i collegamenti ISL tra i siti sono interrotti ed entrambi i siti sono isolati l'uno dall'altro. Come discusso in scenari precedenti, come ad esempio un errore ISL nella rete di gestione e nella rete di storage, le macchine virtuali non sono interessate da un errore ISL completo.

Dopo la partizione degli host ESXi tra i siti, l'agente vSphere ha controlla gli heartbeat del datastore e, in ciascun sito, gli host ESXi locali saranno in grado di aggiornare gli heartbeat del datastore nei rispettivi volumi/LUN di lettura/scrittura. Gli host nel sito A presumono che gli altri host ESXi nel sito B non abbiano avuto esito positivo perché non vi sono heartbeat di rete/datastore. VSphere ha nel sito A tenta di riavviare le

macchine virtuali del sito B, operazione che alla fine ha esito negativo perché i datastore del sito B non saranno accessibili a causa di un guasto all'ISL di storage. Una situazione simile si ripete nel sito B.



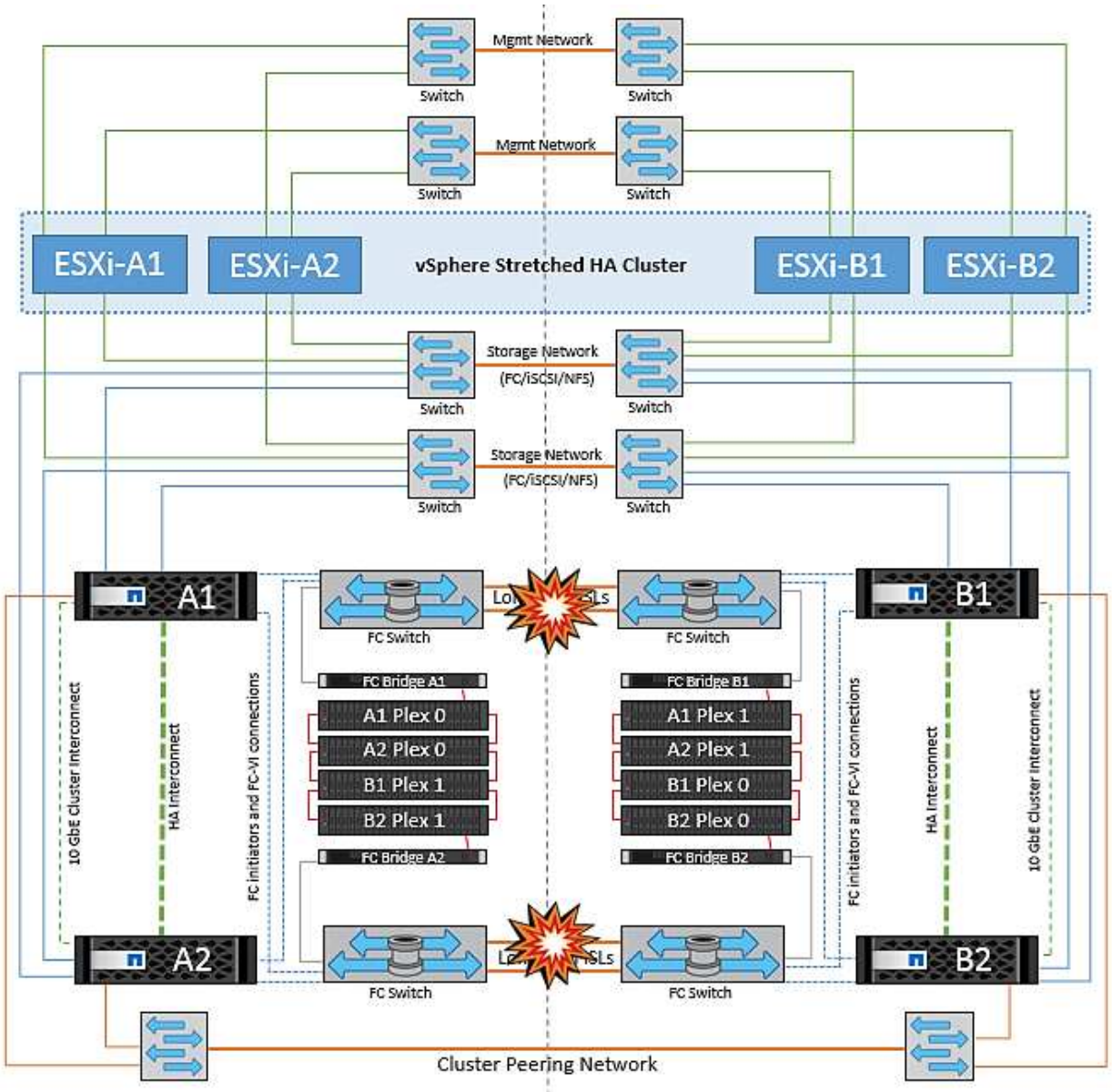
NetApp consiglia di determinare se una macchina virtuale ha violato le regole DRS. Tutte le macchine virtuali in esecuzione da un sito remoto non potranno accedere al datastore, quindi vSphere ha riavvia la macchina virtuale nel sito locale. Una volta che i collegamenti ISL sono tornati in linea, la macchina virtuale in esecuzione nel sito remoto verrà interrotta, poiché non possono esistere due istanze di macchine virtuali in esecuzione con gli stessi indirizzi MAC.



Errore collegamento interswitch su entrambi i fabric in NetApp MetroCluster

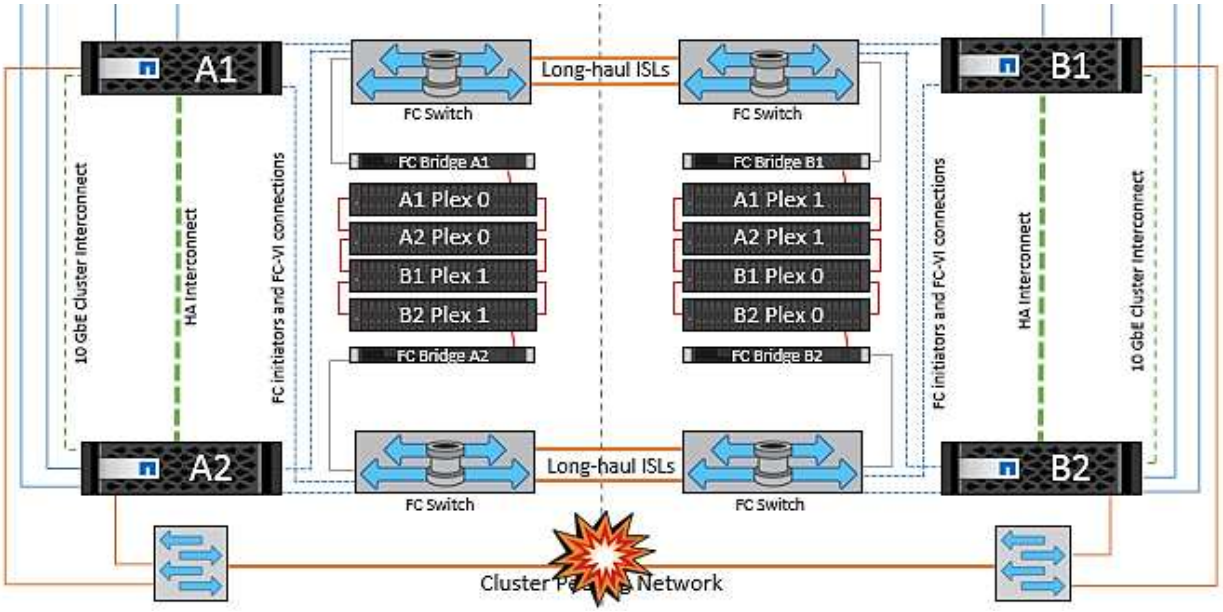
In uno scenario di errore di uno o più ISL, il traffico continua attraverso i collegamenti rimanenti. In caso di errore di tutti gli ISL su entrambi i fabric, in modo da eliminare un collegamento tra i siti per la replica di storage e NVRAM, ciascun controller continuerà a fornire i propri dati locali. Al ripristino di un minimo di un ISL, la risincronizzazione di tutti i plessi avviene automaticamente.

Eventuali scritture che si verificano dopo che tutti gli ISL sono inattivi non verranno mirrorate nell'altro sito. Uno switchover in caso di disastro, mentre la configurazione si trova in questo stato, causerebbe una perdita dei dati non sincronizzati. In questo caso, è necessario un intervento manuale per il ripristino dopo lo switchover. Se è probabile che non saranno disponibili ISL per un periodo prolungato, un amministratore può scegliere di arrestare tutti i servizi dati per evitare il rischio di perdita di dati se occorre eseguire uno switchover in caso di disastro. L'esecuzione di questa azione deve essere valutata rispetto alla probabilità che un evento disastroso richieda lo switchover prima che almeno un ISL diventi disponibile. In alternativa, in caso di errore degli ISL in uno scenario a cascata, un amministratore può attivare uno switchover pianificato verso uno dei siti prima che tutti i collegamenti abbiano avuto esito negativo.



Errore collegamento cluster in peering

In uno scenario di guasto al link del cluster in peering, poiché gli ISL del fabric sono ancora attivi, i servizi dati (letture e scritture) continuano in entrambi i siti verso entrambi i plessi. Eventuali modifiche alla configurazione del cluster, ad esempio l'aggiunta di una nuova SVM, il provisioning di un volume o di una LUN in una SVM esistente, non possono essere propagate all'altro sito. Questi vengono conservati nei volumi di metadati CRS locali e propagati automaticamente all'altro cluster al ripristino del collegamento di cluster sottoposto a peering. Se occorre uno switchover forzato prima del ripristino del link del cluster in peering, le modifiche alla configurazione del cluster in sospeso verranno riprodotte automaticamente dalla copia replicata remota dei volumi di metadati presenti nel sito rimasto nel processo di switchover.



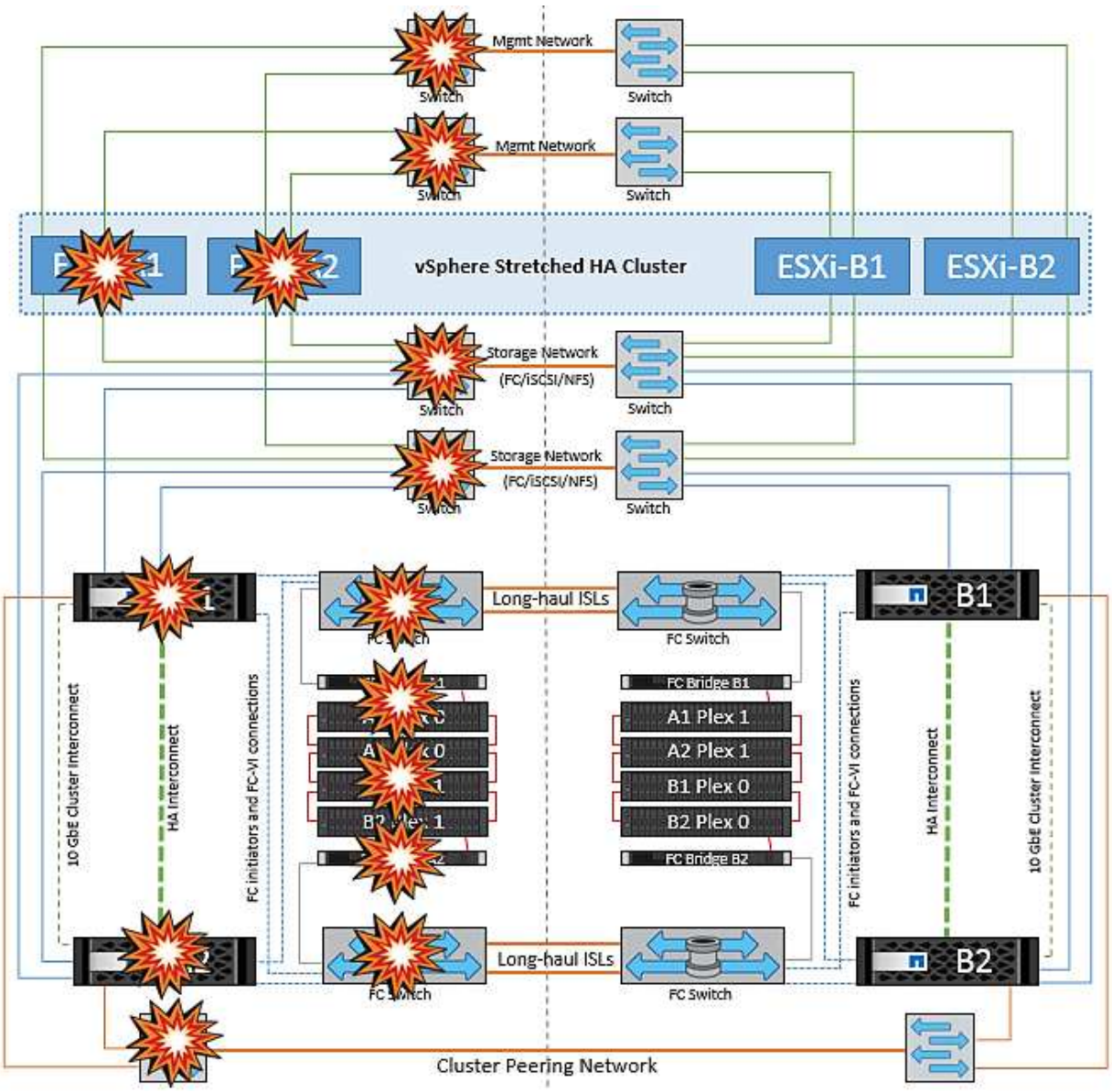
Errore completo del sito

In uno scenario di guasto completo del sito A, gli host ESXi nel sito B non otterranno l'heartbeat di rete dagli host ESXi nel sito A perché non sono attivi. Il master ha nel sito B verificherà che gli heartbeat del datastore non siano presenti, dichiarerà che gli host nel sito A non sono riusciti e tenterà di riavviare le macchine virtuali del sito A nel sito B. Durante questo periodo, l'amministratore dello storage esegue uno switchover per riprendere i servizi dei nodi guasti del sito rimasto e ripristinare i servizi di storage del sito A del sito B. Dopo che i volumi o le LUN del sito A sono disponibili nel sito B, l'agente master ha tenterà di riavviare le macchine virtuali del sito A nel sito B.

Se il tentativo dell'agente master vSphere ha di riavviare una VM (che comporta la registrazione e l'accensione) non riesce, il riavvio viene rieseguito dopo un ritardo. Il ritardo tra i riavvii può essere configurato fino a un massimo di 30 minuti. VSphere ha tenta di riavviare il sistema per un numero massimo di tentativi (sei tentativi per impostazione predefinita).

Nota: il master ha non inizia i tentativi di riavvio fino a quando il placement manager non trova lo spazio di archiviazione adeguato, quindi in caso di un guasto completo del sito, ciò avverrà dopo l'esecuzione dello switchover.

Se il sito A è stato sottoposto a switchover, un guasto successivo di uno dei nodi del sito B sopravvissuto può essere gestito senza problemi attraverso il failover verso il nodo rimasto. In questo caso, il lavoro di quattro nodi viene ora eseguito da un solo nodo. Il ripristino in questo caso consisterebbe nell'esecuzione di un giveback al nodo locale. Quindi, quando il sito A viene ripristinato, viene eseguita un'operazione di switchback per ripristinare il funzionamento regolare della configurazione.


Mgmt Network
Switch
Mgmt Network
Switch
vSphere Stretched HA Cluster
ESXi-B1
ESXi-B2
Storage Network
(FC/iSCSI/NFS)
Switch
Storage Network
(FC/iSCSI/NFS)
Switch
Long-haul ISLs
FC Switch
B1
FC Bridge B1
A1 Plex 1
A2 Plex 1
B1 Plex 0
B2 Plex 0
FC Bridge B2
10 GbE Cluster Interconnect
HA Interconnect
FC initiators and FC-VI connections
FC initiators and FC-VI connections
HA Interconnect
10 GbE Cluster Interconnect
Long-haul ISLs
FC Switch
B2
Cluster Peering Network

Informazioni sul copyright

Informazioni sul marchio commerciale